Development Via Information Self-structuring of Sensorimotor Experience and Interaction

Chrystopher L. Nehaniv[1], Naeem Assif Mirza[1], and Lars Olsson[2]

[1] Adaptive Systems Research Group, University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB, U.K.
{C.L.Nehaniv, N.A.Mirza}@herts.ac.uk
[2] Netemic Ltd., Prospect Studios, 67 Barnes High Street, London SW13 9LE, U.K.
lo@abstractvoid.se

Abstract. We describe how current work in Artificial Intelligence is using rigorous tools from information theory, namely *information distance* and *experience distance* to organize the self-structuring of sensorimotor perception, motor control, and experiential episodes with extended temporal horizon. Experience is operationalized from an embodied agent's own perspective as the flow of values taken by its sensors and effectors (and possibly other internal variables) over a temporal window. Such methods allow an embodied agent to acquire the sensorimotor fields and control structure of its own body, and are being applied to pursue autonomous scaffolded proximal development in the zone between the familiar experience and the unknown.

1 Introduction: Information Self-structuring in Ontogeny

Modern Artificial Intelligence (AI) research has increasingly focused on adaptive, embodied agents with rich sensing capabilities situated in complex environments, that develop in their capabilities over the course of their "lifetimes" (ontogeny) [1, 2]. In our and related research particular attention is paid to the process of autonomous self-structuring in response to a history of self-motivated interaction with a rich environment. The aim is to investigate in artificial agents mechanisms of motivation, learning, development, and temporal awareness with inspiration drawn from biology, psychology, philosophy, engineering, and mathematics.

In this article we review a class of methods for discovering relationships between any and all sensors and actuators that an agent has access to. The methods use the measure of information distance based on Shannon information theory [3] and capture the degree to which the time-varying nature of a variable may be predictable from another. These measures have been used in robots to autonomously discover sensorimotor maps from unknown sensors grounded in interaction with the environment and to discover fundamental control laws for unknown actuators [4, 5] thus gaining mastery over one's own embodiment (which may well be changing). Related classical geometric and statistical methods have also been

M. Lungarella et al. (Eds.): 50 Years of AI, Festschrift, LNAI 4850, pp. 87–98, 2007.

used in simulation to discover sensorimotor relationships [6] and the related structure of space via sensing and acting [7]. Further, the information-theoretic and related methods have been used to characterize behaviour of robots from the robot's perspective [8, 9] and also to measure, geometrically, how one sensorimotor experience differs from another [10]. The self-structuring of the sensory and motor competencies is enabled by the tight coupling of the agent with the environment [11, 2, 5] and the agent can directly base action on its own history of interaction with the environment (including the social environment) to make this possible [12].

2 Information Distance Measures

2.1 Sensors as Information Sources

An agent situated and acting in an environment will have access to many external and internal variables any of which can be modeled as random variables changing over time. These can be thought of as generalized "sensory" inputs, from sources having any character at all (whether sensory, motor, or internal), such as, e.g., registration on sensory surfaces (activations of retinal neurons in vision or cochlear hairs in hearing, readings coming from spatially distributed tactile sensors such as skin and whiskers, etc.), proprioception, motor values sent to effectors, physiological variables, other internal states, etc. Consider one such random variable $\mathcal{X}$ changing with time, taking value $x(t) \in X$, where X is the set of its possible values. Time is taken to be discrete (i.e. t will denote a natural number) and $\mathcal{X}$ takes values in a finite set or "alphabet" $X = \{x_1, \ldots, x_m\}$ of possible values[1].

2.2 Information Distance

For any pair of jointly distributed random variables ("sensors") $\mathcal{X}$ and $\mathcal{Y}$ the *conditional entropy* $H(\mathcal{X}|\mathcal{Y})$ of $\mathcal{X}$ given $\mathcal{Y}$ is the amount of uncertainty (in bits) that remains about the value $\mathcal{X}$ given that the value of $\mathcal{Y}$ is known, and is given by

$$H(\mathcal{X}|\mathcal{Y}) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 \frac{p(x, y)}{p(y)},$$

where $p(x, y)$ is the joint distribution of $\mathcal{X}$ and $\mathcal{Y}$.[2]

The *information distance* between $\mathcal{X}$ and $\mathcal{Y}$ is then defined as

$$d(\mathcal{X}, \mathcal{Y}) = H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y}|\mathcal{X}).$$

[1] The approach generalizes to continuous time and value sets with appropriate changes.

[2] The methods require the assumption of approximate local stationarity of the joint distribution of random variables representing the sensorimotor variables over a temporal window and that this distribution can be estimated closely enough by sampling the sensorimotor variables.

Crutchfield [13] shows that this satisfies the mathematical axioms of equivalence, symmetry and the triangle inequality and so is a *metric*. (See the Appendix for a visual proof of this fact.) Thus d defines a geometric structure on any space of jointly distributed information sources, such as the sensorimotor variables of an embodied agent.

The metric space geometric structure is advantageous as it potentially allows one to exploit the highly developed advanced techniques of mathematical analysis and geometry in the control of behaviour.

2.3 Experience Distance

Given the above definitions we can operationalize an agent's *experience* from time t over a temporal horizon of h time units as $E(t,h) = (\mathcal{X}^1_{t,h}, \ldots, \mathcal{X}^N_{t,h})$ where $\mathcal{X}^1, \ldots, \mathcal{X}^N$ is the set of all sensorimotor (or other) variables available to the agent and each $\mathcal{X}^i_{t,h}$ is the random variable estimated from the values of $\mathcal{X}^i$ over a window of length h beginning at time t $(1 \leq i \leq N)$.

We can then define a metric, the *experience metric* D, on experiences of temporal horizon h as

$$D(E, E') = \sum_{k=1}^{N} d(\mathcal{X}^k_{t,h}, \mathcal{X}^k_{t',h}),$$

or, alternatively, the *cross-modal experience metric* D', as

$$D'(E, E') = \sum_{i=1}^{N} \sum_{j=1}^{N} d(\mathcal{X}^i_{t,h}, \mathcal{X}^j_{t',h}),$$

where $E = E(t,h)$ and $E' = E(t',h)$ are experiences of an agent at time t and t' over horizon h and d is the information distance. That D (and similarly D') is a metric follows from the fact that the metric axioms hold component-wise, since d is a metric.

As experiences are collected, they can be placed in a *metric space of experience* using either of these experience metrics. Experiences generated from similar behaviour as judged from the human observer's perspective generally turn out to be nearby from the robot's perspective in terms of the experience metric in such a space [14].

This operational notion of experience facilitates the application of information-theoretic methods to sensorimotor variables in a way that is directly related to *embodiment*. Such an agent-centric approach already brings these rigorous methods closer to a type of Shannon information that is meaningful for perception and action, however it is possible to go much further and develop a rigorous notion of *relevant information* specific to particular organisms and agents, by relating action, information and utility – see [15].

3 Development of Artificial Cortex: Using Information Theory as a Means for Self-organizing Sensorimotor Structures Grounded in Experience

How can raw, uninterpreted information from unknown sensors come to be used by a developing embodied agent with no prior knowledge of its motor capabilities?

In nature, cognitive structures appear to be organized in the course of evolution and also in the course of development so as to reflect information-theoretic relations arising in the interaction of sensors, actuators, and the environment (including the social environment). That wiring economy for intracortical connectivity of functionally related neural processing structures yields evolutionary fitness has been proposed as a general principle giving rise to topographic structure of cortical maps (see review in [16]) and permitting "extraordinary speed in representing ecologically significant information" [17].

We have applied the information distance metric to develop and reconstruct "retinotopic" and cortex-like sensorimotor maps for robots as they engage in interaction with real-world environments (see Figure 1, and [4, 5] for details). Information distance (rather than mutual information or other measures such as Hamming distance) appears to lead to the best structured cortex-like maps of sensorimotor variables, especially for cross-modal integration [18]. This power might be due to information distance's metric nature, which allows natural geometrization of information sources (which could possibly also give raise to wiring economy), coupled with the capacity of information distance to detect relations between information sources that are informationally, but not necessarily linearly (nor even continuously), correlated.

Even in brain areas much older than the cortex, such as the superior colliculus in mammals (the area homologous to the optic tectum in other vertebrates), cross-modal alignment of visual, auditory, and tactile and somatosensory maps is evident [19]. For instance, in the ferret or barn owl such visual and auditory maps are aligned in this region in proximity to neural pre-motor control circuitry allowing the animal to combine sensory modalities in guiding action, e.g. combining or using either of visual and auditory information in neural maps to guide head movements and eye saccades in orienting toward prey, or, e.g. in reaching in primates; moreover maps are maintained and aligned over the course of development, which may be activity and sensory stimulation dependent – see [20, 21].

In artificial embodied agents, such sensory fields that are constructed on the basis of information distance (see preceding section) methods [4] can then be used to autonomously discover sensorimotor laws (Figure 2), e.g. optical or tactile flow and visually guided movement [5]. The particular embodiment and environment experienced and the changes in it can shape the sensorimotor maps generated in this way, as well as drive their dynamic unfolding and adaptation in ontogeny [4, 5].

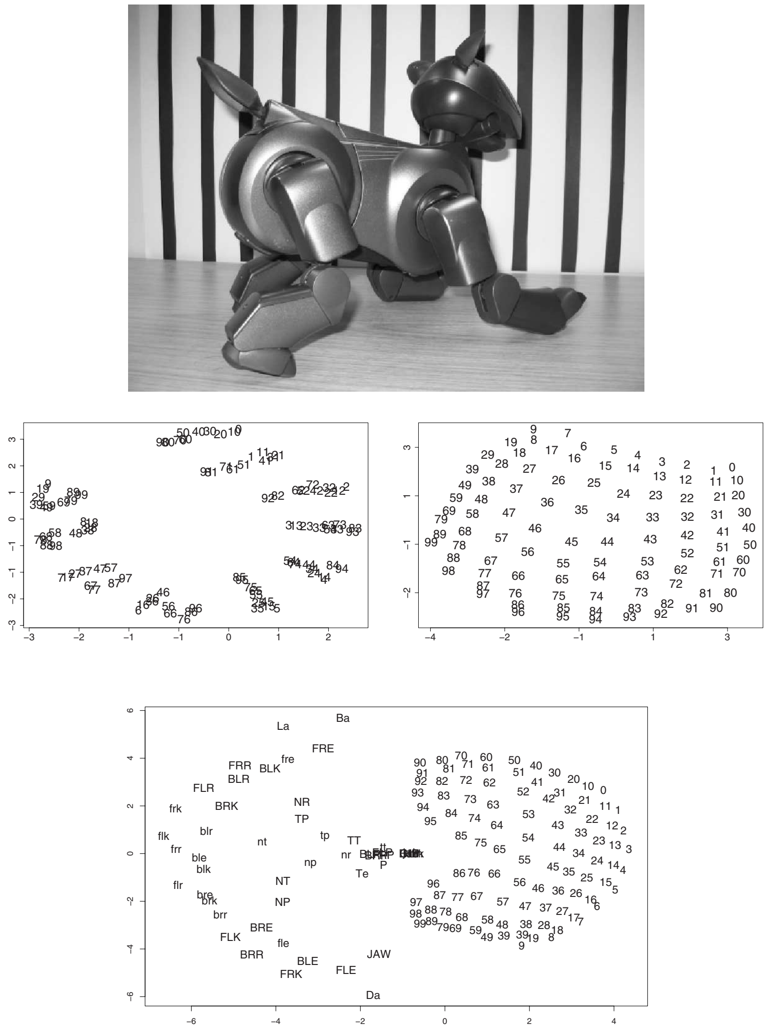

Fig. 1. Sony Aibo in striped environment develops impoverished distinctions between sensors, but further development may allow distinctions to unfold. Top: Robot moving in the striped environment. In the remaining subfigures, points represent individual information sources (sensors or actuators of the robot) plotted using the information distance (and collapsed into two dimensions). Middle left: Sensory organization of the vision sensors (pixels in the visual field) developed in impoverished environment reflects only similar sensory experience of visual receptors from the same columns. Middle right: Sensory organization of vision sensors after moving to richer visual environment reflects their topographical layout. Bottom: Cortex-like "Aibunculus" sensorimotor organization – analogous to the somatosensory "homuncular" cortical maps – recovered based on self-structuring during agent-environment interaction using information distance, discovering visual field (numbered sensors, clustered "retinotopically" and arrayed to the right) and left-right bilateral body symmetry along an anterior-posterior axis.

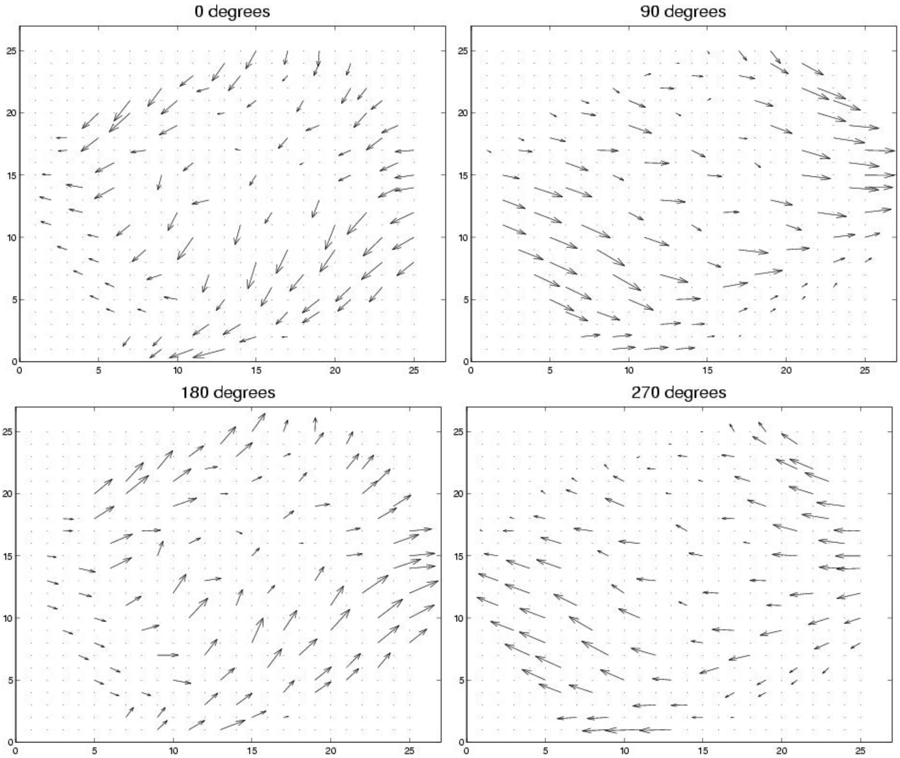

Fig. 2. Sensory fields (in this case a two-dimensional visual field), sensory flows on the fields (the regularities in value shifts due to actions - in this case head movements), and sensorimotor laws (describing the flows resulting from actions) are autonomously derived from experience based on bottom-up information self-structuring and used to control visually guided movement [5]. Figure shows discovered sensorimotor regularities in sensory flows induced by motions of the Aibo's head in various directions, where 0 degrees is up, 180 degrees is down, 270 degrees is right, and 90 degrees is left from the robot perspective.

Alignment of multimodal information sources is demonstrated using the Aibo robot for red, green, blue color channels in vision via *entropy maximization* and *information distance* self-structuring, and this combination of information distance and entropy maximization is shown by far to out-perform other metrics and measures for sensory map construction (see [18] for details). Combining multimodal sensory integration with pre-motor control based on alignment of sensory and body maps is a next natural target for such methods.

4 Temporally Extended Experience and Interaction Histories

How can embodied agents develop in response to extended experiences at various scales in time?

Generally, in AI so far the *temporal horizon* [22] has either been limited to the immediate time-scale (traditional behaviour-based reactivity), short-term modulation (affective computing and robotics), or, if longer term, then has generally remained ungrounded (case-based reasoning or symbolic planning) and not susceptible to developmental change in ontogeny. *Autobiographic agents* dynamically construct and reconstruct their experiences in a process of remembering while actively engaged in interaction with the rest of the world [23].

Using extensions of the information metric to experiential episodes of various temporal horizons (see section 2.3), it is possible to impose geometric order on a robot's temporally extended sensorimotor experiences, at various temporal scales [24]. The structure of these dynamic spaces of experiences provides an agent-centric enactive representation of interaction histories with the environment (including the social environment), grounded in the temporally extended sensorimotor experience and used in generating action [8, 10, 25].

Potentially an agent can act using this dynamically growing, developing space of experiences to return to familiar experiences, predict the effect of continuing on a current behavioural trajectory, and explore at the boundary of what is already mastered (cf. Vygotsky's notion of "zone of proximal development"). By using temporally extended experiences to guide action and interaction, we will have the beginnings of post-reactive robotics and grounded episodic intelligence in artificially constructed enactive agents that grow, develop, and adapt their cognitive structures with a broader temporal horizon.

This possibility is explored in our experiments where a robot uses actions determined by a history of interaction to develop the capability to play the early learning game "peekaboo" of iteratively seeing/revealing and not-seeing/hiding the face with an interaction partner [10, 12]. The architecture uses experience distance (based on information distance) to compare experiences and to place them in a metric space. Actions are chosen based on proximity in this space and motivational value of experience. (See Figure 3, and [12] for details.) Peekaboo not only has inherent simple narrative temporal and rule structure [26], but is also believed to be important in providing scaffolding to young infants in developing social expectations and primary intersubjectivity [27]. By forming expectations and selecting action based on previous temporally extended experiences, the agent is able to develop the capacity to engage in practice of temporally complex skills such as social play, and to re-engage in them when similar experience arises again. It should also be possible to explore at the geometric boundary of already mastered skills and familiar behaviour in the experience metric space, which grows and changes dynamically with the lifelong history of interaction.

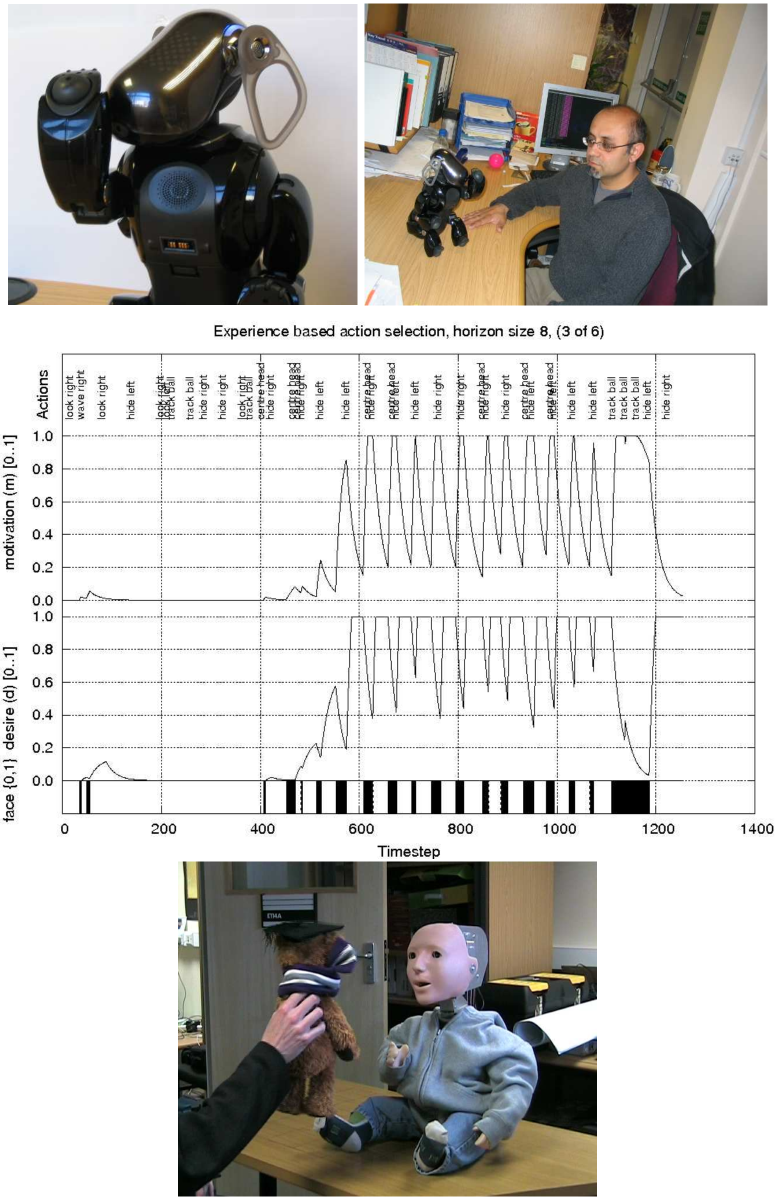

Fig. 3. Top left: Aibo hides face in the autonomous development of the ability to engage in a 'peekaboo' turn-taking game. Top right: Aibo engaging in interaction games with human partner based on interaction history and informationally structured space of experiences. Center: Dynamics of internal variables and actions selected as face is seen or not, with iterations by black/white pattern at bottom of panel indicative of 'peekaboo'-style interaction. Bottom: Interaction games with another platform, the KASPAR child-sized humanoid robot built at University of Hertfordshire.

It may be that using the information and experience distance metrics to organize sensorimotor and episodic experience might capture relations that, in natural organisms, are reflected in the topologies arising from such self-organizing principles as spike time dependent plasticity (cf. [28, 29]) that structure neural connections in development and spatiotemporal sensorimotor pattern learning.

5 Summary and Outlook

Information methods can guide the autonomous organization and structuring of sensorimotor data, lead to the autonomous detection of sensorimotor laws, and underpin the acquisition of sensorimotor control starting with raw uninterpreted sensory data and unknown actuators. Similarly, by extending the methods to encompass sensorimotor flow during particular temporally extended intervals, episodic experience can be operationalized for an embodied system. The geometry of experiences is organized by their information-theoretic structure, and is proposed as a basis for achieving development in robots that grow up, re-engaging in familiar activity, exploring at the boundary of what is already developed, controllable, and mastered. This includes not only sensorimotor experience of static environments, but also interaction histories in dynamic environments involving social interaction or changing embodiment.

Acknowledgments

We are grateful to Daniel Polani, Kerstin Dautenhahn, and René te Boekhorst for collaboration on various aspects of the work surveyed here. This work was partially conducted within the EU Integrated Project RobotCub (Robotic Open-architecture Technology for Cognition, Understanding, and Behaviours), funded by the EC through the E5 Unit (Cognition) of FP6-IST under Contract FP6-004370.

References

[1] Lungarella, M., Metta, G., Pfeifer, R., Sandini, G.: Developmental robotics: A survey. Connection Science 15(4), 151–190 (2003)
[2] Lungarella, M., Sporns, O.: Information self-structuring: Key principles for learning and development. In: ICDL 2005. Proc. of 4th International Conference on Development and Learning, pp. 25–30 (2005)
[3] Shannon, C.E.: A mathematical theory of communication. Bell Systems Technical Journal 27, 379–423, 623–656 (1948)
[4] Olsson, L., Nehaniv, C.L., Polani, D.: Sensory channel grouping and structure from uninterpreted sensor data. In: 2004 NASA/DoD Conference on Evolvable Hardware, Seattle, Washington, USA, June 24-26, 2004, pp. 153–160. IEEE Computer Society Press, Los Alamitos (2004)
[5] Olsson, L., Nehaniv, C.L., Polani, D.: From unknown sensors and actuators to actions grounded in sensorimotor perceptions. Connection Science 18(2), 121–144 (2006)

[6] Pierce, D., Kuipers, B.: Map learning with uninterpreted sensors and effectors. Artificial Intelligence 92, 169–229 (1997)
[7] Philipona, D., O'Regan, J.K., Nadal, J.P.: Is there something out there? Inferring space from sensorimotor dependencies. Neural Computation 15(9) (2003)
[8] Mirza, N.A., Nehaniv, C.L., Dautenhahn, K., te Boekhorst, R.: Using sensory-motor phase-plots to characterise robot-environment interactions. In: CIRA 2005. Proc. of 6th IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp. 581–586. IEEE Computer Society Press, Los Alamitos (2005)
[9] Kaplan, F., Hafner, V.: Mapping the space of skills: An approach for comparing embodied sensorimotor organizations. In: ICDL 2005. Proceedings of the 4th IEEE International Conference on Development and Learning, pp. 129–134. IEEE Computer Society Press, Los Alamitos (2005)
[10] Mirza, N.A., Nehaniv, C.L., Dautenhahn, K., te Boekhorst, R.: Interaction histories: From experience to action and back again. In: ICDL 2006. Proceedings of the 5th IEEE International Conference on Development and Learning, Bloomington, IN, USA, IEEE Computer Society Press, Los Alamitos (2006)
[11] te Boekhorst, R., Lungarella, M., Pfeifer, R.: Dimensionality reduction through sensory-motor coordination. In: Kaynak, O., Alpaydin, E., Öja, E., Xu, L. (eds.) ICANN 2003 and ICONIP 2003. LNCS, vol. 2714, pp. 496–503. Springer, Heidelberg (2003)
[12] Mirza, N.A., Nehaniv, C.L., Dautenhahn, K., te Boekhorst, R.: Grounded sensorimotor interaction histories in an information theoretic metric space for robot ontogeny. Adaptive Behavior 15(2), 167–187 (2007)
[13] Crutchfield, J.: Information and its metric. In: Lam, L., Morris, H.C. (eds.) Nonlinear Structures in Physical Systems - Pattern Formation, Chaos and Waves, pp. 119–130. Springer-Verlag, New York (1990)
[14] Mirza, N.A., Nehaniv, C.L., Dautenhahn, K., te Boekhorst, R.: Using temporal information distance to locate sensorimotor experience in a metric space. In: CEC 2005. Proc. of 2005 IEEE Congress on Evolutionary Computation, Edinburgh, Scotland, UK, 2-5 September 2005, vol. 1, pp. 150–157. IEEE Computer Society Press, Los Alamitos (2005)
[15] Nehaniv, C.L., Polani, D., Olsson, L., Klyubin, A.S.: Information-theoretic modeling of sensory ecology. In: Laubichler, M., Müller, G.B. (eds.) Modeling Biology: Structures, Behavior, Evolution. The Vienna Series in Theoretical Biology, MIT Press, Cambridge (in press)
[16] Chklovskii, D.B., Koulakov, A.A.: Maps in the brain: What can we learn from them? Annu. Rev. Neurosci. 27, 369–392 (2004)
[17] Diamond, M.E., Petersen, R.S., Harris, J.A.: Learning through maps: Functional significance of topographic organization in the primary sensory cortex. J. Neurobiol. 41, 64–68 (1999)
[18] Olsson, L., Nehaniv, C.L., Polani, D.: Measuring informational distances between sensors and sensor integration. In: Artificial Life X, MIT Press, Cambridge (2006)
[19] King, A.J.: Sensory experience and the formation of a computational map of auditory space in the brain. BioEssays 21(11), 900–911 (1999)
[20] Knudsen, E.I., Brainard, M.S.: Creating a unified representation of visual and auditory space in the brain. Annual Review of Neuroscience 18, 19–43 (1995)
[21] King, A.J.: The superior colliculus. Current Biology 14(4), R335–R338 (2004)
[22] Nehaniv, C.L., Dautenhahn, K., Loomes, M.J.: Constructive biology and approaches to temporal grounding in post-reactive robotics. In: McKee, G.T.,

Schenker, P., (eds.) Sensor Fusion and Decentralized Control in Robotics Systems II (September 19-20, Boston, Massachusetts), Proceedings of SPIE. vol. 3839 pp. 156–167 (1999)
[23] Dautenhahn, K., Christaller, T.: Remembering, rehearsal and empathy - towards a social and embodied cognitive psychology for artifacts. In: Ó Nualláin, S., McKevitt, P., Mac Aogáin, E., (eds.) Two Sciences of the Mind: Readings in Cognitive Science and Consciousness. John Benjamins North America Inc. pp. 257–282 (1996)
[24] Nehaniv, C.L.: Sensorimotor experience and its metrics. In: Proc. of 2005 IEEE Congress on Evolutionary Computation, Edinburgh, Scotland, UK, 2-5 September 2005, vol. 1, pp. 142–149. IEEE Computer Society Press, Los Alamitos (2005)
[25] Nehaniv, C.L., Mirza, N.A., Dautenhahn, K., te Boekhorst, R.: Extending the temporal horizon of autonomous robots. In: Murase, K., Sekiyama, K., Kubota, N., Naniwa, T., Sitte, J. (eds.) AMiRE 2005. Proc. of the 3rd International Symposium on Autonomous Minirobots for Research and Edutainment, pp. 389–395. Springer, Heidelberg (2006)
[26] Bruner, J.S., Sherwood, V.: Peekaboo and the learning of rule structures. In: Bruner, J., Jolly, A., Sylva, K. (eds.) Play: Its Role in Development and Evolution, pp. 277–285. Penguin, New York (1975)
[27] Rochat, P., Querido, J.G., Striano, T.: Emerging sensitivity to the timing and structure of protoconversation in early infancy. Developmental Psychology 35(4), 950–957 (1999)
[28] Song, S., Miller, K.D., Abbott, L.F.: Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature Neuroscience 3, 919–926 (2000)
[29] Izhikevich, E.M.: Polychronization: Computation with spikes. Neural Computation 18, 245–282 (2006)

Appendix. Short Proof that Information Distance Satisfies the Axioms for a Metric on Information Sources

We can give a short proof that d is really a metric: Specifically, d is a metric since it satisfies the following axioms for every three (jointly distributed) information sources $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$:

1. $d(\mathcal{X}, \mathcal{Y}) = 0$ if and only if $\mathcal{X}$ and $\mathcal{Y}$ are equivalent (equivalence).
2. $d(\mathcal{X}, \mathcal{Y}) = d(\mathcal{Y}, \mathcal{X})$ (symmetry).
3. $d(\mathcal{X}, \mathcal{Y}) + d(\mathcal{Y}, \mathcal{Z}) \geq d(\mathcal{X}, \mathcal{Z})$ (triangle inequality).

Proof: In the first condition, "equivalent" means "informationally equivalent", i.e. that knowing the value of $\mathcal{X}$ completely determines the value of $\mathcal{Y}$, and vice versa. This can only be the case exactly when both of the conditional entropies are zero. The second condition is trivial from the symmetry of the expression $H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y}|\mathcal{X})$. To see that the triangle inequality holds, draw a "Venn diagram" visualization for the entropies of the three random variables $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ (see Fig. 4). Now the quantity $d(\mathcal{X}, \mathcal{Y})$ corresponds to the "double crescent" region (i.e. excluding the overlap) for $\mathcal{X}$ and $\mathcal{Y}$ representing the sum of their (non-negative) conditional entropies in bits. Now it is obvious that the double crescent for $\mathcal{X}$ and $\mathcal{Y}$ together with the double crescent for $\mathcal{Y}$ and $\mathcal{Z}$ cover the one for the pair $\mathcal{X}$ and $\mathcal{Z}$, and, since for all the variously shaded regions the corresponding entropies are non-negative, it follows from the covering that the inequality holds. □

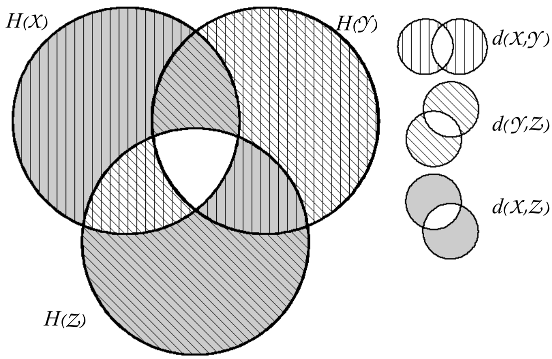

Fig. 4. Visual Proof of the Triangle Inequality for Information Distance. Visualization of the entropies H of three information sources modeled as random variables $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$, with the variously shaded double-crescent regions showing, for each pair of variables, the sum of these conditional entropies, which gives their information distance. Right: Three information distances are visualized as double-crescent regions in the key. Here the left crescent for the information distance $d(\mathcal{X}, \mathcal{Y})$ from $\mathcal{X}$ to $\mathcal{Y}$ represents the conditional entropy $H(\mathcal{X}|\mathcal{Y})$ and the right crescent represents the conditional entropy $H(\mathcal{Y}|\mathcal{X})$; similarly, the other double-crescent regions corresponding to $d(\mathcal{Y}, \mathcal{Z})$ and $d(\mathcal{X}, \mathcal{Z})$ are shown. Left: Venn diagram visualization for the entropies of the three information sources. The triangle inequality holds since the double-crescent region for $d(\mathcal{X}, \mathcal{Z})$ is completely covered by those for $d(\mathcal{X}, \mathcal{Y})$ and $d(\mathcal{Y}, \mathcal{Z})$.