Meaning for Observers and Agents

Chrystopher L. Nehaniv
Interactive Systems Engineering Research Group
Faculty of Engineering & Information Sciences
University of Hertfordshire
College Lane, Hatfield Herts AL10 9AB
United Kingdom
E-mail: `C.L.Nehaniv@herts.ac.uk`
WWW: `http://www.cs.herts.ac.uk/~nehaniv`

Abstract

Claude Shannon formalized the notion of *information* transmission rate and capacity for pre-existing channels. Wittgenstein in his later work insisted that linguistic meaning be defined in terms of *use* in *language games*. C. S. Peirce, the father of semiotics, realized the importance of sign, signified, and interpretant in processes of semiosis. In particular, the connection between sign and signified does not take place in a platonic vacuum but is situated, embodied, embedded, and must be mediated by an interpretant.

We introduce a rigorous mathematical notion of *meaning*, as (1) agent- and observer- perceptible information in interaction games between an agent and its environment or between an agent and other agents, that is (2) *useful* for satisfying homeostatic and other drives, needs, goals or intentions.

With this framework it is possible to address issues of sensor- and actuator- design, origins, evolution, and maintenance for biological and artificial systems. Moreover, correspondences between channels of meaning are exploited by biological entities in predicting the behavior or reading the intent of others, as in predator-prey and social interaction. Social learning, imitation, communication of experience also develop and can be developed on this substrate of shared meaning.

Keywords: information, semiosis, sensors, actuators, automata, interaction games, usage, evolution

1 Introduction

We describe how the study of agents, constructive robotics and biology can benefit from considerations of the origin, design, evolution and maintenance of channels of meaning for various observers and agents. The notion of 'meaning' here will be formalized as a refinement extending the standard Shannon-Weaver measures of information (e.g. [1]).

This formalization relies also on Ludwig Wittgenstein's realization that in considering the 'meaning' of words or 'speech acts' in human language it is essential and sufficient to study their usage [2, 3]. This requires a focus on the particular structural coupling between agents and their environments. Extending this view to channels of information, one is taken beyond the realm of 'language games' into an analysis of 'interaction games'. Rather than restricting to language however we allow a very broad, general notion of signal.

As Peirce showed in his initiation of the science of semiotics, the naive consideration of sign and signified misses an essential third aspect of the semiotic triangle, namely the 'interpretant' which mediates their relations (cf. [4, 5]). In a semiotic view of meaning then, by explicitly taking into account for whom – for which agent or observer – a channel carries useful information, one is led to a refinement of the notion of information [6].

In particular, the identification of agents and observers leads to multiple possible arisings of meaning in a given situation, and sensors (or senses) and actuators constitute candidate targets and sources of information in meaning channels of interaction. Moreover, the consideration of observers and agents leads to multiple semiotic triads, even in the case of only one information channel — and sometimes to none and hence to no meaning (even if entropy is maximal). Signal usefulness to an agent can, as for animals, be grounded in a 'common-currency' of cost and reward that is a basis for neural systems of emotion and learning [7], or via its effect on the probability of reproductive success as in the Darwinian theory of natural selection.

The genesis and continued existence of channels of meaning is not given *a priori* as in the case of information theory. On the contrary, a formal approach to meaning allows one to address just these issues of how channels come into existence and evolve. By identifying sources and targets of information useful to some agents (or attributed as being useful to one by an observer – who may be the agent itself) one has candidate endpoints of a channel of meaning. Over time, classes of channels of useful information need not have sources and targets of static type, and indeed these may evolve in particular agents (as a result of adaptation or learning) or in particular populations (as a result of evolution). For instance, there are multiple cases of evolution of eyes with sensitivity to various light frequency ranges and to various types of changes in a visual field in vertebrate and invertebrate animals. The design and evolution of such meaning channels becomes rigorously addressable within our framework.

Moreover, by considering each source-channel-target structure in a situated context, 'first-person structures' are identified, i.e. particular instances of agents interacting with their environments via their sensors and actuators, as they experience the world. Mappings or correspondences between such structures – via algebraic homomorphisms – help explain how the experience of one agent can be meaningful for an observer or for an interaction partner. The study of such mappings, relations and interaction is the study of 'second-person structures', which are correspondences between the structure and experience of agents. Thus human observers often attribute emotional and intentional states by mapping their own types of first-person experience to other agents, such as other people, infants, and animals, or even to simple mechanisms as in the case of Braitenberg's vehicles [8, 9]. Agent-environment interactions such as building which may seem 'meaningless' or merely reactive for an individual agent such as with termite stigmery [10] may from the viewpoint of a colony as a unit of selection serve to maximize its reproductive success and be quite meaningful.

Seond-person mapping opens the possibilities for the evolution of language and other forms of interactive communication by taking advantage of inter-agent correspondences in first-person structures and channels of meaning. Building of robotic agents along constructive lines (e.g. [11, 12]), and considerations of interaction between agents, such as observation, language, play, imitation, prey-seeking, predator-avoidance, mate attraction, imitation, and communication, may thus benefit from exploiting a rigorous theory of meaning channels, their capacity, origins, maintenance, construction, design, and evolution.

2 Meaning for an Agent

2.1 Usefulness of Channels: Evolution of Sensors, Actuators, and Meaning

Particular instances of information, signals or symbols may have no meaning for an agent. This is obvious if one considers that they might fail to be perceptible or have any effect on the agent. The fact that a biological agent perceives or is otherwise influenced reveals a potential for a signal to be important to the agent. In particular, the fact that a biological agent has sensors that can detect any of a particular class of signal, such as eyes that respond to light in a certain range of wavelengths, are evidence that it is useful to do so. Thus some channels of information may not have any meaning for an agent, yet for others sensors may evolve to tune into certain information sources in a manner that is adaptive, i.e. tending to increase fitness, survival, reproductive success. Similarly, channels of action, via actuators, for acting on the world and for producing signals, may evolve for reasons of fitness. These channels of sensing and acting for an animal are means for it to interact with its world, seek out food, prey, escape from predators, change its internal state in response to environmental changes, recognize and court potential mates, etc. The information in these channels is meaningful for the agent since it is useful for the agent in achieving *goals* of homeostasis, survival, and reproduction; if it is capable of having intentions and behavioral goals, meaningful channels of interaction must have evolved to support these as well.

Biological evolution may act on higher level units of fitness than the individual, for example a colony of ants, termites, or bees with one or few reproductive individuals is a unit of selection on which evolution acts. Indeed, differentiated multicellular life arises from a cooperative population of cells, reproductive units, which have given up some control over their reproduction, by differentiation into soma and germ lines, for the benefit of being part of a successful higher level unit of evolution (an animal or higher plant) where copies of their genes are more likely to persist through time than if they were to reproduce at a maximal rate (cancer) [13, 14, 15]. Even these lower level units, the cells themselves, evolved from symbiotic associations of bacteria to form eukaryotic cells with their mitochondria and chloroplasts: The eukaryotic cell is a "garden of bacteria" (Margulis [16]).

For a higher level unit of selection, channels of perception and action useful for its goals of survival and reproduction are also meaningful. Information channels for any agent that help it attain its goals are meaningful. Having 'goals' is not taken to imply nor to exclude the possibility of intentionality behind them. The goals can be merely states or classes of states that agent under appropriate conditions has a tendency to move to (e.g. as is the case with body-temperature regulation, adequate oxygen and glucose levels in blood, and other homeostatic tendencies). Goal states are also those which by virtue of the organismal dynamics as modulated by hormonal control, drives, or emotion (state-change in response to reinforcing stimuli) lead to behaviors to obtain or avoid a stimulus [17, 7]; other goal states can possibly be those resulting from processes of deliberation and planning (e.g. in humans). Thus for all kinds of agents (biological, synthetic, software, lower or higher level unit of evolution), whether information is *useful* in attaining any of their goals (in this technical and general sense) is the criterion by which meaningfulness of information and action can be judged.

3 Automata Models of Agents

Regard the set X of possible states of an agent as given.[1] Transitions between the states can be induced by receiving external signals or internally in response to change in or decay of certain variables (e.g. internal clocks (central pattern generators, or circadian rhythym regulators), oxygen- or glucose- level in blood, etc.). Denote the triggers of such transitions by the event set Σ. Some of the elements of Σ may be complex, e.g. an image array of retinal activation, or the vector of inputs with components from all senses. We shall suppose that the next state of the agent depends on current state $x \in X$ and a transition event $s \in \Sigma$ which triggers that transition.

Moreover, we shall suppose that actions by the agent can be determined from its current state. That is, being in a state x may induce a behavior of the agent such output, signalling, use of actuators, intention movements, etc. This execution of behavior may itself induce a change of state. Thus initiation of the behavior is itself to be regarded as a member or component of a member of some event $s \in \Sigma$. Thus the agent is represented as an automaton (X, Σ).

This notion of automaton that allows internal arising of transitions is similar to the augmented finite-state automata of Brooks [19] used in subsumption architectures where such automata are combined in layers such that some of them may modulate the event sets of others. However, our automata need not be finite-state.

Algebraic automata theory allows one to compute the transformation semigroup (X, S) for the automaton (X, Σ) by identifying (collapsing the set of) all sequences of events in Σ which induce the same mapping from the state set X to itself. The transformation semigroup (especially, if finite) can be decomposed using the Krohn-Rhodes Theorem and its generalizations into irreducible components, and its computational power and complexity can then be studied [20, 21].

3.1 Agents, Observers, and Interaction

Several agents can of course interact with each and with the worlds around them. The events in the environment may serve as triggering events to some, all or none of these agents. The behavior of one agent (resulting in signals, displays, actions, etc.) given in the specification of the automaton describing that agent may in turn serve as a triggering event for transitions in other agents. In such cases, the environment or respectively the first agent is a *source*, the event is a *signal*, and the agent in whom an event is triggered is a *target*. Conversely, states (with corresponding actuator activations and signals produced) within the first agent may also act on the environment.

Here we have three types of 'channels of meaning' with source-target pairs: environment-agent, agent-agent, agent-environment. There may be several such channels between agents, such that each agent is a target for a channel the other is a source of. The emergent behavior of the system of agents and environment in such a case is referred to as an *interaction game*. Following and generalizing the ideas of Wittgenstein we say *meaning of the signals can be and can only be defined in terms in their usage in interaction games.*

It is to be expected that evolution will act to ensure that sensor and actuator channels used in recurring types interaction games will over generations to some degree be optimized in order to better achieve each agent's goals of survival and reproduction. Since many interaction partners may be present (e.g. both predators on the agent and potential mates), and signals may be perceived by many interaction partners, selection pressures will result in trade-offs in the signals and behaviors the agent exhibits.

---

[1] Alternatively, following James P. Crutchfield [18], construct X as equivalence classes of states which make future observations conditionally independent from the past, i.e. the probability distribution over all possible sequences of future events depends on only the current state. In this case, the transitions between states are given by single observations.

3.2 Transmission, Reception, and Entropy

The information transmission capacity of channels of meaning can be defined along the lines of Shannon information theory [1]. The *transmission capacity* of the channel is

$$C = \lim_{T \to \infty} N(T)/T,$$

where $N(T)$ is the number of possible signal sequences produced by the agent during T transitions. (Note that we are using the events giving transitions of the agent as a 'clock' for the time with respect to the agent.) The *measure of information content* for each signal produced by the agent in a given state x is given as its entropy,

$$H_x = -\sum_{i=1}^{n} p_i \log p_i,$$

where the sum is taken over all possible signals, with n signals possible in state x, and $p_i \in [0, 1]$ is the probability of signal i in this state. Then the *entropy of the agent as a source for this channel* is

$$H = \sum_x f_x H_x,$$

where f_x is the average frequency (occurrences per transition) of state x.

Evolutionary considerations suggest that agent actuators should evolve in such a manner that the optimal transmission rate C/H of signals per transition event of the source agent for this channel is maximal subject trade-offs for redundancy in dealing with noise as well as trade-offs resulting from other observers being able to access the channel (e.g. predator hearing a mating call or seeing an intraspecies display). Thus, an evolved agent will have signals and behavior as meaningful possible (transmission rate at close to the optimal rate), i.e. as useful as possible in achieving its goals (homeostasis, reproduction, etc. - see above).

From the target's (or observer's) viewpoint the received signals may not be the same set as for the sender. The observer may partition the signal stimuli differently, or only react to some aspect of the signal (e.g. be color blind), or detect more structure in the signal than the sender can actively control (i.e. the signal may carry extra information the sender cannot perceive which may vary with its state, e.g. facial expressions and posture in humans during speech.) An evolved agent is expected to extract as much meaning (in this formal sense) as possible from its perceptions.

Because one has different agents (or just the environment) at one end of the channel, there may be this asymmetry between signals sent and the transition events (elements of Σ) received. Thus, in addition to a *transmission capacity* for the channel, one has also a *reception capacity* for the channel, in which one replaces the word 'source' by 'target', 'produced' by 'received', and 'signal' by 'transition event'. From evolutionary consideration, one concludes that the sensors should evolve so that reception rate of the channel should be maximal with trade-offs resulting from the cost of building sensors and the adaptive advantage that they provide. The capacity for transmission is optimized by evolving or designing actuators and behaviors in a source agent and reception capacity with sensors in a target agent.

We have analyzed transmission and reception by agents, but if at one end of the channel are non-living aspects of the environment rather than any biological agent, then we do not expect that this part of the environment will evolve subject to Darwinian evolution or optimize its transmission or reception. Nevertheless we do not expect that it will remain static over evolutionary time.

3.3 The Example of Squid Displays

Many cephalopods (squids, cuttlefish, octopuses) have evolved fascinating visual body patterns and signalling displays used in intraspecies interaction (courtship behavior, aggression between males), and interspecies interaction: predator avoidance and hunting prey (camouflage, confusing 'protean' displays, 'passing cloud', inking). Many aspects of their complex body patterning are under extremely fast neural control, making their whole bodies into display devices, and portions of their bodies that are visible to different observers can be controlled independently (e.g. aggressive display by a male on the side facing another male with concurrent courtship display toward a female on the other side, or inking perceived by an aereal predator while hunting with display directed toward prey as the squid continues to hunt under the ink). See [22, 23].

Much of this signalling can go on independently of body postures and movement. The displays and patterning of these animals evolved in response to a complex environment where there are many potential observers of various species. This evolution must have been shaped by trade-offs between the reception-transmission rates in the meaningful channels where this signalling occurs and the pressures from predators, obtaining prey, and interaction with conspecifics. Martin Moynihan has suggested that the number of rare displays is constrained by possible difficulty in their interpretation [23]; this conclusion is independently supported by our considerations here: by the fact that rare displays have low probability by definition, the information content of a squid's displays is reduced if the proportion of rare displays increases. However, in the case of rare signals, in order to optimize transmission, it is to be expected that they are encoded to have duration over more transitions, whereas common signals are likely to have short duration.

So far we have not addressed the relative cost of signals. Rare signals could invoke more expensive mechanisms (such as inking in squids). Furthermore, for a more realistic picture of the optimization of signalling it will be necessary to take account of the costs and benefits such as cost of producing a particular signal, cost and likelihood of detection, and benefit from the communication.

4 Discussion and Directions

Warren Weaver [1, pp. 24–28] in his introduction to Shannon's mathematical theory of communication expressed the hope that information theory could be developed into a theory of meaning, and we have indicated how this can be done. Our notion of meaning depends on agents and observers (sources and targets of signal channels) and is grounded on the usage and usefulness of sensory information and actuator activity for achieving an agent's 'goals'. In the case of biological agents, these goals include homeostasis, survival and reproduction, and may include others such as actual plans and intentions. The notion of meaning, unlike that of information, thus depends on agents. In contrast to what Weaver envisioned [1, p. 26], there is no notion of external 'semantic noise' in a channel corresponding to 'noise' in the sense of information theory, since semantics (meaning) is found only with agents and observers. Noise can still act on signals in a channel of meaning, but the semantics arises only at endpoints of channels where agents reside. Meaning is not external to agents, but only makes sense with respect to their sensing and acting in the world via interaction games.

The notion of automata (and transformation semigroup) mappings can be used to study the relations of the particular agent to others. These mappings (homomorphisms) are structure-preserving correspondences between different agents with their sensor-actuator channels. More generally than homomorphisms, relational morphisms are mappings which provide broad correspondences in which sets of states and events may be related (e.g. [24]).

Much work remains to be done in the study of shared channels of meaning, and in the projection of these channels between agents. For example, an observer of an interaction game between two other agents gives raise to a meaning by virtue of its observation. It may attribute meaning to the interaction in a way that preserves structure: as Braitenberg [8] and Pfeifer [9] have discussed, emotional states such as 'love', 'aggression' and 'fear' are attributed by human observers to reactive robot models engaged in taxis and other simple behaviors. This is indicative of a tendency to project one's own experience of meaning to other agents (humans, animals, robots, etc.) and allows us to make sense of their actions. By attributing an internal state and assuming that they act to achieve their goals in a manner similar to the way we do, we anthropomorphize the other and in many cases are then able to predict its behavior. This capacity for projection (mapping) of the experience of meaning may indeed be very important in the animal mind (predicting the behavior of prey, predators, or conspecifics) and social interaction and intelligence [25, 17, 26, 27]. One could formally approach this problem via an analysis of channels of meaning between agents and structure-preserving mappings that establish the correspondences in the experiences of various agents. Further problems of temporal grounding could also be addressed along rigorous mathematical lines using algebraic automata theory, and algebras of time and history; and indeed communication of histories (narrative) could represent particularly interesting types of signals in channels of meaning for temporally grounded agents (See [28, 25]).

The conclusions concerning optimal transmission and reception rates for meaning in evolved agents which were derived above on theoretical grounds could be validated or rejected by the use of evolutionary computation simulations — where it should be easy to manipulate or eliminate the trade-offs (sensor cost, predator attention, etc.) that could not be controlled in a natural or laboratory setting involving living organisms.

The theory introduced here can also be extended to continuous channels of meaning as in classical information theory.

References

[1] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* University of Illinois Press, 1963.

[2] L. Wittgenstein, *The Blue and Brown Books.* Harper & Brothers, 1958.

[3] L. Wittgenstein, *Philosophical Investigations,*

(Philosophische Untersuchungen), German with English translation by G. E. M. Anscombe, 1964. Basil Blackwell, Oxford, reprinted 3rd edition, 1968.

[4] C. S. Peirce, *Collected Papers*, vol. 2: Elements of Logic. Harvard, 1965.

[5] J. Goguen, "An introduction to algebraic semiotics, with application to user interface design," in *Computation for Metaphors, Analogy, and Agents* (C. L. Nehaniv, ed.), vol. 1562, Lecture Notes in Artificial Intelligence, pp. 242–291, Springer Verlag, 1999.

[6] C. L. Nehaniv, "The second person - meaning and metaphor," in *Computation for Metaphors, Analogy, and Agents*, vol. 1562, Lecture Notes in Artificial Intelligence, pp. 380–388, Springer Verlag, 1999.

[7] E. T. Rolls, *The Brain and Emotion.* Oxford, 1999.

[8] V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology.* MIT Press, 1986.

[9] R. Pfeifer, "The fungus eater approach to emotion: a view from artificial intelligence," *Cognitive Studies*, vol. 1, pp. 42–57, 1994.

[10] P. P. Grassé, "La reconstruction du nid et les coordinations inter-individuelles chez *bellicositermes natalensis* et *cubitermes* sp. la theorie de la stigmergie: Essai d'interpretation des termites constructeurs," *Ins. Soc.*, vol. 6, pp. 41–48, 1959.

[11] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal Robotics and Automation*, vol. RA-2, pp. 14–23, April 1986.

[12] B. Webb, "Using robots to model animals: A cricket test," *Robotics and Autonomous Systems*, vol. 16, pp. 117–134, 1995.

[13] L. W. Buss, *The Evolution of Individuality.* Princeton University Press, 1987.

[14] R. E. Michod and D. Roze, "Cooperation and conflict in the evolution of individuality. III. transitions in the unit of fitness," in *Mathematical and Computational Biology: Computational Morphogenesis, Hierarchical Complexity, and Digital Evolution* (C. L. Nehaniv, ed.), vol. 26 (Lectures on Mathematics in the Life Sciences), pp. 47–91, American Mathematical Society, 1999.

[15] J. Maynard Smith and E. Szathmáry, *The Major Transitions in Evolution.* W. H. Freeman, 1995.

[16] L. Margulis, *Symbiosis in Cell Evolution.* W. H. Freeman, 1981.

[17] C. L. Nehaniv, "Constructive biology of emotion systems: First and second person methods in grounding adaptation in a biological and social world," submitted for journal publication.

[18] J. P. Crutchfield, "Observing complexity and the complexity of observation," in *Inside versus Outside* (H. Atmanspacher, ed.), pp. 234–272, Springer Verlag, 1994.

[19] R. A. Brooks, "The Behavior Language: User's guide," *MIT Artificial Intelligence Lab Memo 1227*, April 1990.

[20] K. B. Krohn, B. R. Tilson, and J. L. Rhodes, "Lectures on finite semigroups": Chapters 1, 5, 7-9, of *Algebraic Theory of Machines, Languages, and Semigroups* (M. Arbib, ed.), Academic Press, 1968.

[21] C. L. Nehaniv and J. L. Rhodes, "On the manner in which biological complexity may grow," in *Mathematical and Computational Biology: Computational Morphogenesis, Hierarchical Complexity, and Digital Evolution*, vol. 26 (Lectures on Mathematics in the Life Sciences), pp. 93–101, American Mathematical Society, 1999.

[22] R. T. Hanlon and J. B. Messenger, *Cephalopod Behaviour.* Cambridge University Press, 1996.

[23] M. Moynihan, *Communication and Non-Communication in Cephalopods.* Indiana University Press, 1985.

[24] C. L. Nehaniv, "From relation to emulation: The covering lemma for transformation semigroups," *Journal of Pure & Applied Algebra*, vol. 107, pp. 75–87, 1996.

[25] C. L. Nehaniv and K. Dautenhahn, "Embodiment and memories — algebras of time and history for autobiographic agents," in *Cybernetics and Systems '98 : Proc. 14th European Meeting on Cybernetics and Systems Research (Vienna, Austria – 14-17 April 1998)* (R. Trappl, ed.), vol. 2, pp. 651–656, Austrian Society for Cybernetic Studies, 1998.

[26] R. W. Byrne and A. Whiten, *Machiavellian Intelligence.* Clarendon Press, 1988.

[27] K. Dautenhahn, "I could be you - the phenomenological dimension of social understanding," *Cybernetics and Systems Journal, Special Issue on Epistemological Aspects of Embodied AI*, vol. 28(5), pp. 417–453, 1997.

[28] C. L. Nehaniv, "What's your story? - irreversibility, algebra, autobiographic agents," in *Socially Intelligent Agents (Proc. AAAI Fall Symposium - November 1997, MIT)*, vol. FS-97-02, pp. 150–153, American Association for Artificial Intelligence, 1997.