Processing networks for the classification of motion — a proposal

Christoph Flores

Millstätter Str. 17
D-28359 Bremen, Germany
flores@informatik.uni-bremen.de

Abstract. Spoken-of motion is usually conceptualised as a change of existential states (an entity moves from A to B in order to change its potentiality for further action), while perceived motion is usually only conceptualised as a change of position. As a means of technically tracking motions, we can think of processing-networks tied to both, sensors and read-out devices, that use a process orientated internal evaluation of incoming data. Properly chosen, they eliminate the difference between detecting a shift of state and detecting a shift of position. A proposal for the layout and use of processing elements in such a network is presented — as an early general construction idea with an invitation to ask questions about its usability.

Keywords. Processing-Networks, State Change during Motion, Motion Classification

Motion as tracked in language is usually a shift of states, while motion as tracked in perception is usually a shift of position. Jill can be *seen* to follow a certain trajectory when crossing the hallway, but she is *described* as leaving the kitchen and walking over to the dining hall. We cannot, in a purely perceptional perspective, "see" her moving out of the kitchen and into the dining hall. Such concepts as "kitchen" and "walking over" which are needed for the cognitively enriched form of seeing are therefore, in our current machine implementations, treated as additional to perceptional motion tracking. They might for example be taken as derivative, in the sense that they can be looked up in a knowledge base that holds information about regions for the kitchen, given in spatial coordinates, regions for the dining hall, and some reachability information that indicates whether there is access from one area to the other. In such an implementation, a computer, after seeing the trajectory, can find out whether a labelling of getting from A to B ("kitchen" to "dining hall") is appropriate or not. It can proceed from *seeing* to *seeing-as* (or "interpreted seeing"). In a certain sense we can conceive of those motions that we have linguistic labels for, as purposeful instantiations of motion patterns.

Dagstuhl Seminar Proceedings 05491
Spatial Cognition: Specialization and Integration
http://drops.dagstuhl.de/opus/volltexte/2007/1031

We shall, in this paper, not discuss motions as static instantiations of intentions. Certainly we often linguistically refer to ongoing change of position as exhibiting constant instantiation of a motion pattern ("Betty is dancing in the hall," "The robot is moving towards the door"). Yet an entity starting or ending such a motion goes through a shift of its existential state, and it is this aspect that is in the focus of our distinction between interpreted and uninterpreted motion. When we uninterpretedly "see" a motion, we can (as a reaction to a change in sensor signals) implicitly know that a state shift took place but we do not "see" what kind of shift this was, unless we add interpretation.

If we intend to computationally enact such interpretation, we might be tempted to find a common format of representation for our sensor data that allows for elements of bundling (bundling of occupied regions in space) and indication of transition (transition in space, and in state space). The conceptualisation of a representation format, in this context, is a traditional candidate, but a representation might not be fully appropriate. Because it is always representation-process-pairs that establish a grounded processing of environmental information, we might seek for a rather dynamic and process-bound way of tying a machine's states to its environment's state, that entails both, detection and interpretation of the environment's state.

We might base our approach on some speculative insights. I propose that in our implementation work we try to fulfil some of the ontological criteria that we find both, in the cosmological concepts of the late WHITEHEAD [1], and in the (unwritten) semiotic ontology of PEIRCE (cf. [2, 3]).[1] According to Whitehead, the basic elements which this world is made of are "actual entities" (sometimes also called "actual events"). They pick up information from a relevant environment (a process which is determined by what data is available), then individually and freely process them to some internal emphasis, and then, at the end of their life-cycle, hand it over to successors while they disappear. Data transition (external process) is determined, while data evaluation (internal process) is free. For Peirce, a potential infinity in data transition that conserves some qualities of the original events is constitutive for a semiosis. This conservative element allows for freedom in the transition process. The aspired-at infinity of the sign process is real just as a mathematical limes is real. Despite never being reached, it is real by being aspired at.

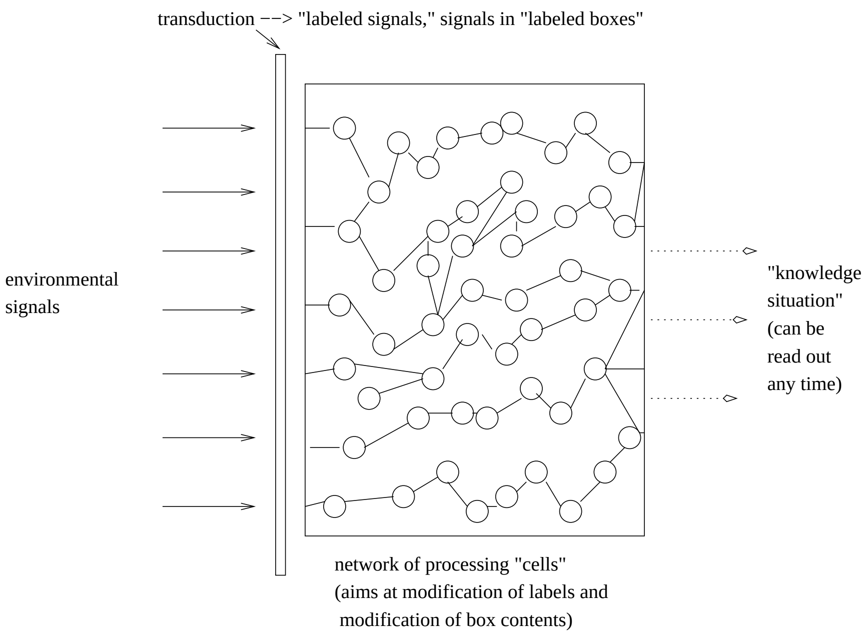
I suggest that we equip our machine with a high number of perceptual units that serve the rôle of data collectors, with internal processing possibilities that aim at increasing some aspects of the incoming data while decreasing or eliminating other aspects of them. After processing, they distribute their results to successors. (This marks an end of a life-cycle.) They can then technically be reused in further such events of streaming data through the system, at a later point in processing time. Some of the perceptual units are connected to input sensors, most of them are connected to each other, and some of them are connected to

[1] Ontology is used here in the traditional sense, as related to what constitutes being as opposed to not being.

output devices. The difference to neural network implementations is that we do not so much care about transition weights for activations from one cell to another, but that we allow some internal processing that changes incoming data. Instead of spreading activation patterns, we spread (processed) data.

We apply judgement algorithms that operate within the linked cells. They strengthen or weaken (or eliminate) input information, and do some of the transition calculations towards the output. We can think of having a hierarchical structure that would entail that those processing units at the beginnings of the chains have other input categories than those towards the ends of the chains, so that they can be used for data intake into the network. Alternatively we can think of having unique entities and a transition layer at the input side of the network. In either case, output qualities of the network are made up by combinatorial use of cells' data, so that output is actually emergent to the grouping of individual output data (so that always some cells form nexūs and constitute a processing aim).

To make a proposal of how our network could be sketched, here's a diagram:



I invite the reader to discuss whether this kind of "machinery in the machine" sounds appealing. Among the questions that need to be answered, and should be discussed, is: "What kind of action patterns (or motion patterns) of a machine that internally makes use of such a nexus of cells can we conceive of?" and "What type of motion classification and action planning capacities can we see as emergent of this potential behaviour?".

In order to try to motivate the approach of chain-processing for sensor signals, I'd like to add one more thought: Our language conceptualises certain movement patterns as possible, others as not (an elephant cannot fly, while a bird can). That does not necessarily imply, though, that all our linguistically coded knowledge of moving objects is depending on a word-net–like tagging mechanism for assignment of possible-actions to objects. But we can conceive of it to be so, nevertheless, at least if we plan to adopt our capacities in machines. A moving object is an entity tagged with possibilities and non-possibilities, indicating prominently what can be known about possible change and non-change in movement, among other things. In order to test for an appropriate labelling, laws of motion mechanics could be applied.

The idea now is that instead of implementing all knowledge about possible change and non-change of state in some explicit mechanism with a "labelling capacity," we put initial data acquisition for motion signals on the sensors, spread this data in the network and use the network's response for ascriptions of classificatory quasi-labels: habit-like reactions to input patterns. So we count on adaptation to input to develop in the processing chain — a "knowledge" or "feeling" for what patterns input-signals have. It is an idea of conserving and strengthening inherent information in data by streaming them through evaluation processes of enforcement or selection or omission.

The evaluation processes are internal to the processing cells. Because they provide the dynamics of the system, they must of course be chosen rich enough to do some useful work — if we want the processing system to exhibit something like "intentions" towards certain processing results, this has to emerge out of our implementations of cell level evaluation processes. Intentionality might ideally be inherent to the evaluation processes, although it is not explicitly coded at implementation time.

I'm very interested to discuss with the reader the potentials that we can find in this general idea, and to intensify speculations aimed at cell implementations of appropriate processes for motion classification.

References

1. Alfred North Whitehead. *Process and Reality. An Essay in Cosmology.* The Free Press, New York, corrected edition, 1978.
2. Charles S. Peirce. *Naturordnung und Zeichenprozeß. Schriften über Semiotik und Naturphilosophie.* Suhrkamp, Frankfurt am Main, 1991.
3. Helmut Pape. *Erfahrung und Wirklichkeit als Zeichenprozeß: Charles S. Peirces Entwurf einer spekulativen Grammatik des Seins.* Suhrkamp, Frankfurt am Main, 1989.