parsing and generation (considered from a competence view) are non-deterministic processes, i.e., without any further (non-grammatical) information both processes have inherent degrees of freedom; cf. also Shieber (1993). Furthermore, the mentioned properties (reversibility, implicit search space, and modularity) are also important in the context of the "recursion-only hypothesis" discussed by Hauser, Chomsky & Fitch (2002), who claim that recursion (i.e., providing the capacity to generate an infinite range of expressions from a finite set of elements) is the only uniquely human component of the faculty of language.

10 However, at least from a language usage perspective, concrete language utterances seem to be deterministic, i.e., at some point some decisions are made. What is the nature of these decisions, if they are not grammatical? At least two possibilities can be considered: either the decisions are based on preferences (which are learned through past experience) or through control, i.e., explicit strategies that are used to interpret the results of other processes in order to provide feedback. I consider preferences as "un-intelligent" in the sense that they are merely applied (blindly) and control-strategies as "intelligent" because they are applied purposely. Of course, both aspects are somewhat integrated, i.e., language processing is both preference-directed and controlled. It seems that the interleaved approach and the anticipatory drive at least have important properties that classify them as an explicit control-strategy.

11 As already said at the beginning, the proposed computational model is mainly rooted in mathematics and computational linguistics and does not claim any cognitive "realism". However, the realization of the underlying ideas (i.e., grammar reversibility, uniform parsing and generation, self-monitoring) on such a technical algorithmic level requires fine-grained details. Furthermore, the idea of the interleaved approach of parsing and generation is also strongly motivated by the assumption that complex anticipation feedback loops are necessary for the modelling of highly self-adaptive natural language systems. And as such, it might be of interest for the further outline of the model of the anticipatory drive proposed by Butz, especially concerning the aspects of language and modularity. Clearly, the computational model in its current form is realized on a high symbolic level. Probably, it is too high to integrate it directly on a neuronal level. Seen as such, it could be of scientific interest to explore a) how to integrate sub-symbolic approaches into such computational models as I have outlined, and b) how to integrate such complex symbolic interactions into the model of the anticipatory drive.

Objectifying the Subjective Self

An Account From a Synthetic Robotics Approach

Jun Tani
RIKEN Brain Science Institute (Japan)
<tani@brain.riken.jp>

1 The current essay by Butz brilliantly illustrates a constructivist account for one of the essential problems in psychology and cognitive science: that of how the subjectively perceived self can be objectified. His theory stands on so-called "anticipatory behavior" (§4), which is considered to play an important role in learning behavioral causalities in the environment that forms the inner reality during development. Butz links two distinct pathways in brains – a dorsal one and a ventral one (§59) – in which the former constructs a body space based on proprioceptional encoding of body postures and the latter does so for visual categorizations of objects. His interesting argument is that bidirectional interactions between these two pathways initiate the objectification process of the subjective self, especially during tool usage as described in Iriki's studies (§61). During the use of each familiarized tool, a distinct sensory-motor structure appears in seamless coupling with parietal neuronal activities that in turn subjectify the tool usage within the body space. On the other hand, when the tool is detached from the body, the tool that was once subjectified within the body space, is now objectified via visual categorization within the object-centered coordinate system in the ventral pathway. Finally, he postulates that this process of objectification of formerly-subjectified tool usage might lead to the objectification of the subjective self (§89).

2 Being impressed by Butz's psychological account for the process of objectification of the subjective self, I would like to postulate, from my expertise in synthetic neuro-robotics studies, possible neuro-dynamic mechanisms that account for his psychological theorem.

3 Before considering the actual mechanisms, I would like to start by discussing differences between notions of self and self-consciousness. A particular concern is that the state of self-consciousness in the reflective stage might not occur just by being able to anticipate motor-caused sensory feedback. Instead, the self might become consciously aware only when a prediction goes wrong, generating errors.

4 This interpretation of self-consciousness may be supported by Heidegger's (1962) example of the hammer, which is well-known in phenomenology. For a carpenter, when everything is going smoothly, the carpenter and the hammer function as a single unit. But, when something goes wrong with the carpenter's hammering or with the hammer, then the independent existences of the subject (the carpenter) and the object (the hammer) are noticed by the carpenter. Here, the carpenter becomes self-conscious, in the same ways that he or she becomes conscious of the world becoming problematic when things just do not match expectations.

5 Tani (1998) reconstructed this phenomena in his neuro-robotics experiments with emphasis of the cognitive roles of regression for learning from the past and prediction of the future. In this experiment, a mobile robot with vision learned to predict the next landmarks it would encounter while it explored the environment by utilizing a dynamic neural network as a forward model. After a certain period of exploration, the robot became predictable by achieving coherence between prediction by the internal forward dynamics and the environmental sensory feedback given when such coherence broke down intermittently, generating prediction error. In the incoherence phase, the internal process is stressed, with a search for better internal parameter values in order to reduce prediction error; while in the coherent phase everything goes smoothly and automatically without the need to update the parameters. In analogy to the Heidegger's example of hammering, it can be said that the robot became self-conscious only in the incoherent phase.

In other words, the self can exist in the coherent phase, but it can be consciously aware only in the incoherent phase. It is further argued that self-consciousness, which appeared only intermittently in Tani's robot, may well represent the momentary self described by William James (1950). Gallagher (2000) regards this type of self as a minimal self, which is only a momentary, subjective experience of self, which may correspond to the reflexive state of self-consciousness introduced in the current essay. Gallagher (2000) wrote that the minimal self can be developed to a narrative self that is constituted with a past and a future in the various stories that we tell about ourselves. This self-referential nature of narrative self seems to correspond well to the reflective stage of self consciousness in the current essay by Butz. It is also noted that this development from the reflexive stage to the reflective one can be related to the transition from the pre-empirical level to the objective time level in Husserl's theorem on immanent time (Husserl 1964), as will be illustrated later.

6 Now, I will propose possible neuronal mechanisms for extending the reflexive stage of self-consciousness to the reflective one in Butz's terminologies. Although the apparently difficult part is how to objectify the subjective inner reality of sensory-motor experiences, this can be modeled by taking two different neuronal representation approaches: namely, those of local representation and of distributed representation. In the local representation approach, each distinct sensory-motor structure experienced can be embedded in its corresponding local forward model module through winner-take-all (WTA) type competitions with other modules (Wolpert & Kawato 1998; Tani & Nolfi 1998; see Figure 1a).

7 The competition proceeds with a gating mechanism associated with each module. If a particular forward model module is good at predicting the coming sensory flow while generating less error compared to others, the gate associated with this module tends to open more, while others do so less, in the WTA manner. The winning module is entitled to more learning and generation of more prediction outputs for the current inputs. As a result of the competitive learning, distinct sensory-motor primitives, in terms of forward models, are self-organized into corresponding local modules. After this learning, the original sen-

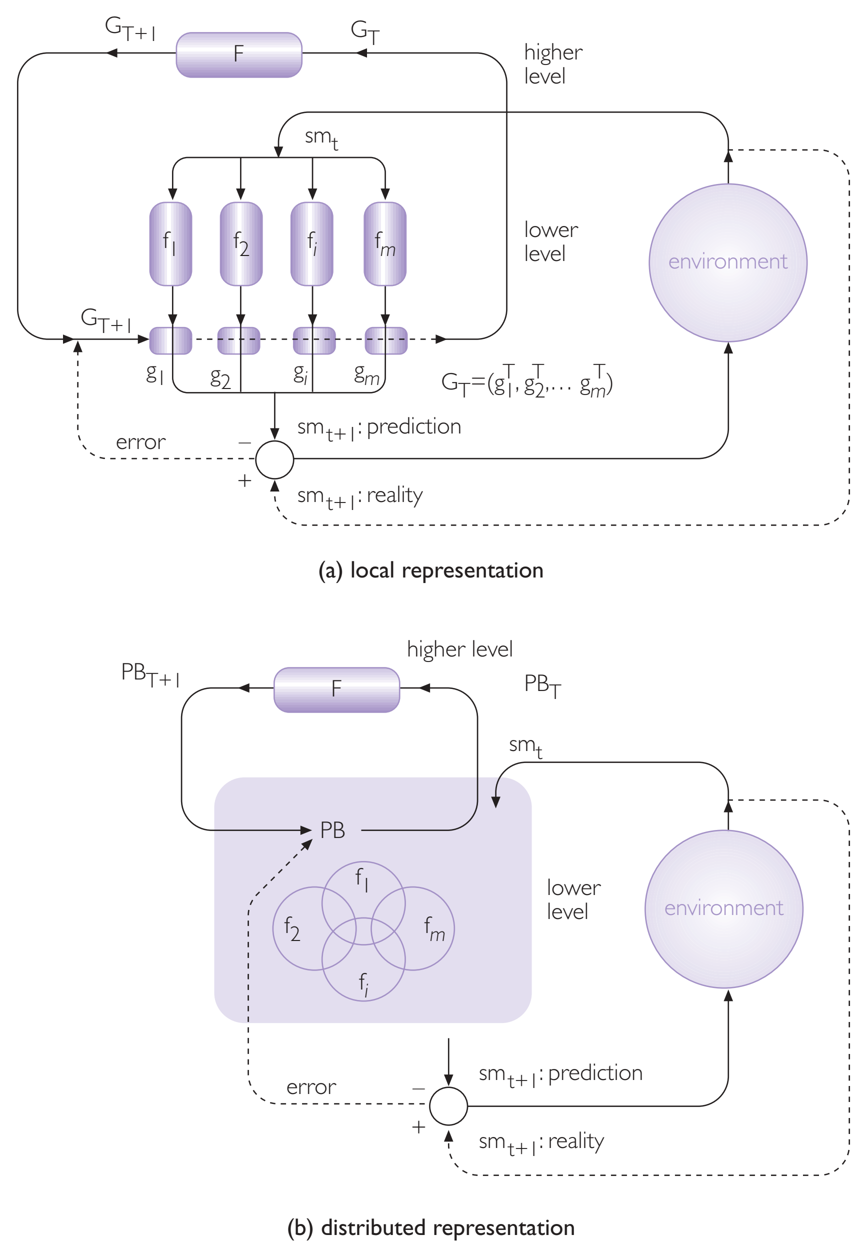

Figure 1: A local representation scheme of the gated modular networks model is shown in (a). In the lower level each *i*-th forward model depicted by f_i competes, using winner-take-all dynamics, to predict the next sensory-motor state sm_{t+1}, while the higher level forward model F predicts which gate will open in the next time step in G_{T+1}. (b) shows a distributed representation scheme of the RNNPB, where multiple forward models, $f_1, ...fm$, are distributionally represented in a single RNN with the associated PB vector. The PB vector value is adapted such that the prediction error for the next sensory-motor state sm_{t+1} is minimized, while the higher level forward model F predicts the next PB vector value, PB_{T+1}.

sory-motor flow is segmented into sequences of reusable primitives by the gate opening mechanism (Tani & Nolfi 1998). A higher level forward model is now introduced to the system (Tani & Nolfi 1998). The higher level forward model learns to predict the gate opening sequences by observing them. Therefore, the original sensory-motor flow is reconstructed in terms of sequences of pre-acquired primitives. Our essential claim here is that subjective experiences of sensory-motor flow are objectified by referencing them with their corresponding module IDs which are manipulable at the higher level. It is noted that prediction error is an essential drive to articulating continuous sensory-motor flow into objectified primitives. In other words, subjective experience is segmented by momentary self-accompanying incoherence into a sequence of consciously retrievable events that constructs the narrative self. At this very moment, the objective time level might appear from the pre-empirical level (see more precise descriptions in Tani 2004).

8 A distributed representation scheme is now introduced where similar but more psychologically plausible explanations can be made (see Figure 1b). Tani and colleagues (Tani, Ito & Sugita 2004) have proposed a neural network model, the so-called "recurrent neural network" with parametric biases (RNNPB), that can learn multiple forward dynamics models in a distributed way within a single recurrent neural network (RNN). In this model, an RNN is associated with additional units, the so-called "parametric biases" (PB). The PB play the role of bifurcation parameters for the forward dynamics realized by the RNN. By modulating the values of the PB vector, the forward dynamics generates diverse dynamic patterns by going through successive bifurcations. The learning in RNNPB is considered as a process of determining an optimal synaptic weights matrix that embeds all the target dynamic patterns and a set of PB vectors specific to each of the target dynamic patterns. As the result of learning, a mapping between the PB vector and the dynamic patterns is self-organized. In the RNNPB, it is considered that each distinct sensory-motor structure is objectified by its corresponding PB vector value. If a higher level RNN is introduced in order to learn sequences of PB vector shifting, a corresponding switching of sensory-motor structures can be obtained at the lower level that seems to be analogous to the gate switching shown in the local representation scheme.

9 However, there is a distinct advantage to the generalization capability of the RNNPB, which originates from its distributed representation characteristics. In the mapping of PB, if the hamming distance between two PB vectors is short, dynamic patterns generated from these two PB vectors become similar. In other cases, they become different from each other. In this manner, the PB mapping can provide a continuous functional space with generalization, while the gating networks cannot attain such a generalization capability because their functional space is partitioned discretely by a finite set of local modular functions.

10 Such generalization characteristics have been demonstrated by an RNNPB-implemented humanoid robotics experiment in manipulating different shapes of objects (Nishide et al. 2008). As has been said in the current essay by Butz (§51), different objects entail different sensory-motor structures. Nishide et al. (2008) trained two types of mappings where one was a PB mapping to the motor trajectories of the robot arms and the other was from visual images of objects to the PB vector. As a result of simultaneous training of these two mappings, when the robot sees one of the trained objects, the visual mapping generates a corresponding PB vector, which turns out to generate the correct motor trajectory for manipulating the object. When the robot was asked to manipulate a novel object for which the visual feature is between two of the pre-trained objects, the motor trajectory was adequately generated as interpolation between two motor trajectories trained for these objects. This generalization capability for novel objects results from the fact that the objectified entities are still represented in the low-dimensional metric space of the PB. Furthermore, when the robot arm was guided by researchers to move using pre-trained motor patterns, the corresponding mental imagery of the visual object was generated because mapping from proprioception to vision through the PB is established by means of the inverse computation. This might be a possible implementation of Butz's idea (§59) of bidirectional mapping between the dorsal processing, specialized for bodyspace encoding, and the ventral processing for object identification during tool use.

11 In the current commentary, two possible neuronal mechanisms have been proposed to account for the psychological pathways of the development of self-consciousness from its reflexive stage to the reflective one, as proposed by Butz. Although both the local representation scheme and the distributed one are shown to be capable of mapping from subjective sensory-motor experiences to objectified entities, the latter might provide a more psychologically-plausible mechanism because the objectified entities still remain in a metric space. Because these objectified entities that appear in the PB space are not like the arbitrary shapes of tokens (Harnad 1990) but preserve metricity, they could have inherently natural interfaces with the sensory-motor reality in the shared metric space.

Anticipation of Motor Acts
Good for Sportsmen, Bad for Thinkers

John G. Taylor
King's College London
<john.g.taylor@kcl.ac.uk>

1 This paper is full of stimulating and creative ideas. It posits that an anticipatory drive is what guides the development in the brain of a set of internal motor models, specifically a set of inverse and forward models. Through these models becoming increasingly complex, a conscious self develops. This is a simple and important thesis, if true. But is it? As my title suggests, it may be so for sportsmen, with their emphasis on ever more refined motor responses. However, those of a more cerebral nature may find themselves burdened by all those coupled internal motor models and not able to think as clearly as they would like. This is not to say that prediction isn't a useful property to possess, both for finance (especially now) and in one's general living patterns. But the question I wish to consider is: What sort of predictive model can lead to thinking?

2 There is a further difficulty with this paper: it promised an answer to "Why consciousness?" Consciousness is claimed to arise from the increasing plethora of internal