Bridging the gap between optimal feedback control and sensorimotor neurophysiology

Emanuel Todorov, Ben Dongsung Huh
University of California San Diego

Optimality principles of sensorimotor function have been very successful in explaining behavior. But their implementation in neural networks, and the ability of the resulting networks to explain neurophysiological data, remain open questions. Here we describe our ongoing effort to address these questions. The main idea is conceptually straightforward: replace the optimal estimator and controller at the heart of optimal feedback control models (**Fig 1a**) with a recurrent neural network (**Fig 1b**), which has the same inputs and outputs and is optimized for the same behavioral cost. Such use of neural networks is unusual because the network outputs do not affect the cost directly, but only indirectly – through their effects of the musculo-skeletal plant. An indirect cost does not specify in any detail what representation the network should use. The emergent neural representation is thus a genuine prediction, obtained from first principles.

Our **first model** is an isometric version of the Kakei et al (99) task, where monkeys made center-out wrist movements in different forearm postures. When forearm posture varied from pronation to supination, muscle preferred directions (PDs) rotated by about 70deg while PDs in primary motor cortex (M1) rotated only about half as much. M1 neurons also showed postural gain changes. The latter are sufficient to rescue our prior model of direct cortical control of muscle activation (Todorov 2000). Nevertheless, this dataset raises an important question: if M1 neurons control muscles directly, why do they behave differently from muscles? The new model presented here suggests the following answer: the redundancy in the M1-to-muscle mapping allows an infinite variety of M1 representations, and the one which happens to be optimal under noise is a mixed muscle-movement representation. The model is a stochastic recurrent network (**Fig 2a**) which receives two distributed inputs (forearm posture and extrinsic target direction – sample input patterns are shown as green and magenta), and is trained to settle into any population output that can be linearly mapped into appropriate muscle activations (two agonist-antagonist pairs of idealized wrist muscles). The muscle activations after training were stochastic, due to 20% signal-dependent noise in the synaptic transmission mechanism, but clustered around the correct activations (solid lines in **Fig 2b**). Shown are results for two different targets (red and green dots), reached in all forearm postures. The recurrent network learned to shift the neuron preferred directions as a function of posture. This was accomplished by a novel mechanism using asymmetric local connectivity (**Fig 2c**). In this mechanism each neuron votes for some preferred direction, but at the same time pushes the population hill of activity away from its preferred direction, in a way consistent with its preferred posture (opposite postural preferences are indicated with orange and black). Overall, the posture-dependent shift of preferred directions was about half of the shift in muscles – in agreement with data. When the network connectivity was constrained to better reflect M1 neuroanatomy, the distribution of shifts became bimodal (**Fig 2d**).

Our **second model** involves a detailed 2-link 6-muscle arm, trained on reaching movements with arbitrary start and end points. To enable such training, we had to modify the backpropagation-through-time algorithm significantly. The new algorithm propagates cost gradients through mixtures of neural networks and nonlinear dynamical systems (such as arm models). This makes it possible, for the first time, to efficiently train recurrent neural networks in the context of stochastic optimal control. Figure 1 illustrates the behavior of a network with 20 neurons, all receiving sensory inputs signaling muscle lengths, velocities, tensions, and target coordinates. All neurons can send outputs to all muscles. The input weights, output weights, and recurrent weights are optimized with the modified backpropagation-through-time algorithm. **Fig 3a** shows hand paths for 10cm movements to 12 targets. **Fig 3b** shows how the activity of each neuron (around movement onset) varies with target direction. Note the cosine pattern. **Fig 3c** shows the distribution of preferred directions. When the initial hand position was rotated around the shoulder, the neuron preferred directions rotated in the same way (not shown). We still have to address the issue of local minima, analyze the network behavior more fully in light of experimental data, and try to understand the mechanism which the network uses to accomplish this task. But it is already clear that we have a very powerful tool for translating high-level optimality principles into detailed neurophysiological predictions.

Fig 1a

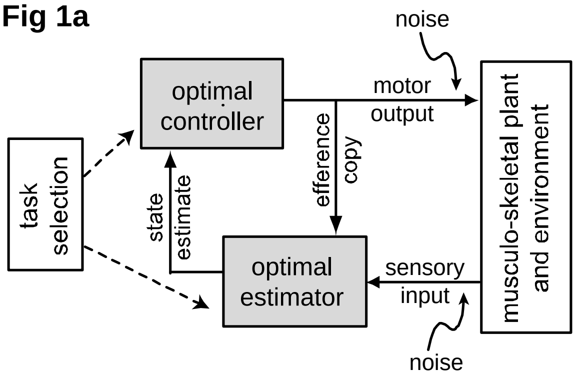

Fig 1b

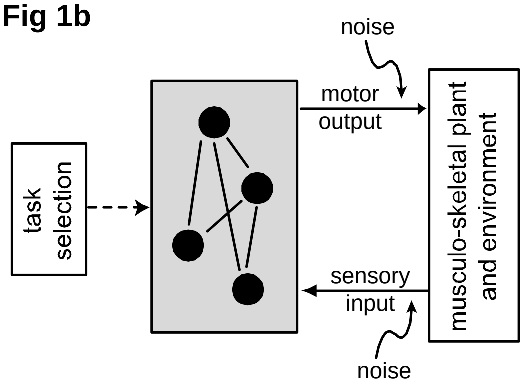

Fig 2a

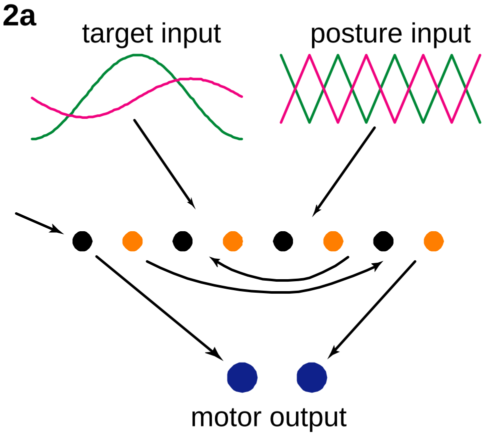

Fig 2c

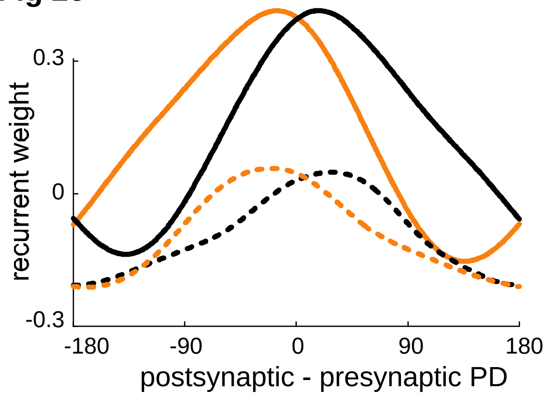

Fig 2b

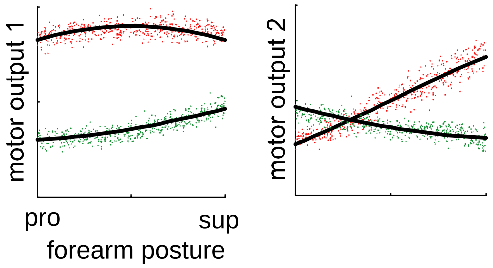

Fig 2d

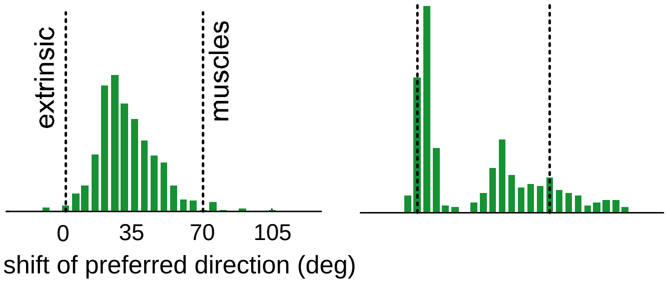

Fig 3a

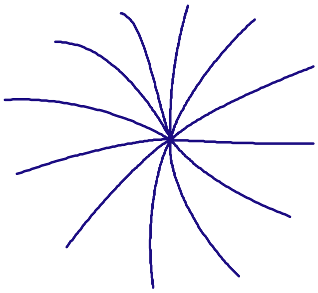

Fig 3b

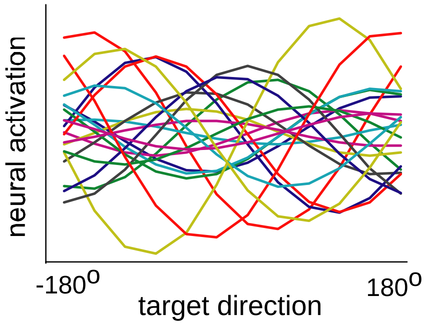

Fig 3c