Infinite Horizon Optimal Control Framework for Goal Directed Movements

Dongsung Huh[1,2], Emanuel Todorov[3], Terrence Sejnowski[1,2]
Salk Institute[1], UCSD[2], University of Washington[3]

Reaching is the simplest and the most studied type of discrete movements, i.e. movements that have finite durations. For obvious reasons, most optimal control models of reaching were formulated in the finite-horizon framework with a pre-specified terminal time. However, we have discovered significant limitations of such formulations, and investigated the alternative, an infinite-horizon formulation.

The earlier finite-horizon models [1,2,3] minimized certain energy functions (e.g. jerk, torque change, variance, or control cost) subject to explicit state-constraints (at target position, zero speed, and zero acceleration) at terminal time. But, these models could only describe open-loop control, insufficient for describing the full complexity of biological feedback control.

More recent finite-horizon models [4] now include feedback, but they have other shortcomings. In these models, the terminal-state-constraints have been replaced with terminal-cost terms that penalize the difference between the desired target-state and the actual state at terminal time. The increased number of cost terms implies more free parameters to adjust (relative weights between the cost terms), but there exists neither a principle for setting the weights nor a sensible interpretation of them. Furthermore, these models have an internal flaw – their optimal trajectories never reach the target-state but leave non-zero error at terminal time (even in deterministic cases!). We can easily show that zero-error trajectories are indeed suboptimal in these formulations.

Here, we propose an alternative formulation in the infinite-horizon framework. Our model is natively feedback-based and it does not include any terminal time constraints (well, there is no terminal time). Instead, the cost function has state-dependent terms (called accuracy cost) as well as an energy term. We found that in order to describe reaching movements, the accuracy cost must have a saturating shape. A similar accuracy cost was corroborated in a recent human psychophysics experiment [5]. Our cost function has two parameters – the width of accuracy cost δ, and a *tempo* parameter γ. Our model also assumes control-dependent (multiplicative) noise in limb dynamics (similar to the minimum variance model [3]), and therefore another parameter – noise coefficient σ.

During an initial acceleration phase, the state-variance increases due to the control-dependent noise, but it gets regulated by feedback control. By the time the target-state is reached, the state-variance is reduced to the desired level of accuracy, δ (Fig 1C). This prediction of time-varying state-variance may be experimentally tested.

In our formulation, movement duration is not a pre-determined constant, but a function of movement amplitude (initial distance from target). This function also depends on the parameters δ, γ, and σ. A large γ increases overall speed of movements and lowers movement duration (Fig 2A,C). However, the model also predicts an upper limit of the overall speed, determined by the noise coefficient σ. Any faster movements would generate more noise than the feedback system can sufficiently reduce, and becomes suboptimal. In the limit $\gamma \rightarrow \infty$, the model precisely produces the speed-accuracy tradeoff curve known as Fitts' law (Fig 3).

Our model explains how movement duration gets determined by both motor-noise (σ) and motivation level (γ) in a unified manner. This explanation is consistent with recent findings, which show that the slow movements of Parkinsonian patients are due to decreased motivation, not increased motor noise [6]. The authors also points out the motivation level (γ) may be related to the tonic level of dopamine, as observed in Parkinson's diseases and Tourette's syndrome patients.

Our model easily generalizes to account for a series of reaching movements by allowing the target-location to be time-varying: $x^*(t)$ (and for that matter, $\gamma(t)$ and $\delta(t)$). Then, we model $x^*(t)$ as a random process (called jump process) which consists of a series of discrete target-jump events. If the average jump rate is low enough to allow enough time intervals for reaching, minimizing the same cost function generates a series of reaching movements.

So what does it mean for a reaching task to be an infinite-horizon control problem? To see this, consider a simplified brain which consists of a high-level area (HLA), and a low-level control area (LLA, e.g. spinal cord) in a feed-forward structure. From the perspective of the HLA, each reaching movement may be a separate, discrete motor task. It sends out a well planned command signal (e.g. $x^*(t)$) to LLA. From the perspective of LLA, however, the task is a continuous one – translating $x^*(t)$ into appropriate time-varying muscle activations. This portion of the task is indeed infinite-horizon in nature, and it is this portion that our model aims to describe. The same insight may apply to other discrete movements as well.

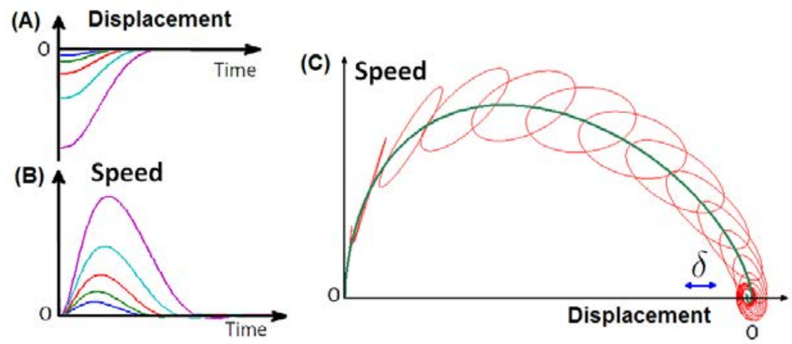

Figure1: (A,B) Average trajectories and speed profiles for movements for various amplitudes (δ, and γ are fixed). (C) State-space representation of an average trajectory (solid-line) and covariance (ellipses). The scale-bar represents the target-width (δ).

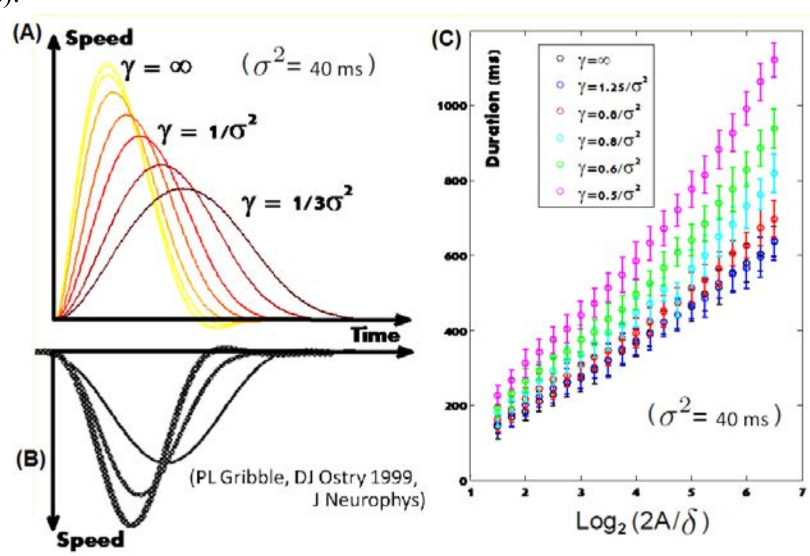

Figure2: (A) Predicted speed-profiles for various γ (movement amplitude and δ are fixed). Faster movements (larger γ) have skewed speed-profile and with some overshooting.
(B) Data from PL Gribble, DJ Ostry (1999) J of Neurophys Vol. 82 No. 5 November 1999, pp. 2310-2326.
(C) Movement duration for various γ. Duration is plotted as a function of movement amplitude divided by δ. (Open-circles = average duration, Error-bar=standard deviation).

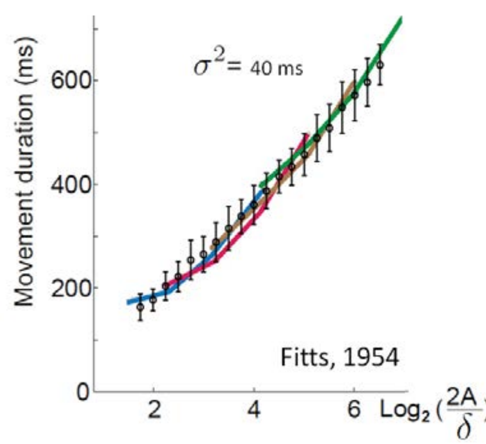

Figure3: Speed-accuracy tradeoff curve. (Open-circles & error-bars) Model prediction for γ→∞ (from Fig 2C). (Solid-lines) Experimental data from PM Fitts (1954). *J of Exp Psychology*, volume 47, number 6, June 1954, pp. 381–391

[1] T. Flash and N. Hogan. *J Neuro*, 5(7):1688, 1985.
[2] Y. Uno, M. Kawato, and R. Suzuki. *Bio cybernetics*, 61(2):89–101, 1989.
[3] C. Harris and D. Wolpert. *Nature*, 394(6695):780–784, 1998.
[4] E. Todorov and M.I. Jordan. *Nature* neuro, 5(11):1226–1235, 2002.
[5] K.P. Koerding and D.M. Wolpert. *PNAS*, 101(26):9839, 2004.
[6] P. Mazzoni, A. Hristova, and J.W. Krakauer. *J Neuro*, 27(27):7105, 2007.