Computable Trust in Human Instruction

Brenna D. Argall[1,3] and **Todd D. Murphey**[2]
[1]Departments of Electrical Engineering & Computer Science and Physical Medicine & Rehabilitation
[2]Department of Mechanical Engineering
Northwestern University, Evanston, IL 60208, USA
[3]Rehabilitation Institute of Chicago, Chicago, IL 60211, USA

Introduction

For the majority of real-world cyber-physical systems, including the domain of humans interacting with robots, intuitive procedures for designing controllers are needed. While one such approach is to teach a robot via demonstration, the control behaviors produced by this data-driven technique are unable to be verified for feasibility or stability. In contrast, sophisticated stability analysis is possible for control behaviors derived via optimal control, but such formulations are rarely intuitive and often require substantial expertise to use.

We propose an approach that creates a *synergy* between *intuitive design interfaces* for a physical system and the *formal verification* that control provides. In particular, our approach derives control behaviors using optimal control while simultaneously engaging a human operator to provide *physical guidance* for adaptation via *corrective demonstration.* A fundamental technical challenge lies in the fact that the operator may well destabilize a system that operates in the physical world, subject to dynamics and sources of uncertainty; moreover, the risk to the system changes from one operator to another. Controllers developed under our approach are *verified for stability and robustness*, and a *formal measure of trust* in the teacher is used to decide whether to cede control to the teacher during physical correction.

For many systems, physically demonstrating motion need not be a particularly challenging task. When a system is suitably quasi-static, simply utilizing some form of impedance control may well allow for physical demonstration—and because there is no risk of destabilization, only the most rudimentary knowledge of the system kinematics is needed. Within *Learning from Demonstration* (Argall et al. 2009), typically an intuitive understanding by the teacher of the kinematics and dynamics of the learner holds—either because s/he has a good understanding of the system platform, or a good idea of how to perform the task with their own body. In contrast, our work specifically allows for differences in dynamics and controls between the learner and teacher. We additionally aim to quantify limits on what the teacher is required to know about the learner for demonstrated corrections to be effective.

We consider the fact that for cyber-physical and human-robot systems, an understanding of *each by the other* is of crucial importance. That is, not only does the *human operator* need to understand the automated system in order to provide good shaping guidance and sound control input, but likewise the *automated system* needs to understand the quality limitations on the guidance and controls provided by the operator. Moreover, this circular dependence has been shown to be quite problematic in that the more complex the software system is, the harder it is for the human operator to interact with it in a meaningful and predictable manner. The proposed automated system therefore will reason explicitly about its trust in the operator's instruction.

We present a mathematical formalism for a computable measure of trust in human instruction. We frame this formalism within the more general scope of cyber-physical systems, for which a human teacher instructing a robot learner is a specific instance. *For physical systems like a robot, reasoning about the stability and utility of instruction is essential not only for good performance but also for safety.*

A Framework for Mutually Controlled Motion

Key to our vision for how mutually controlled motion can be achieved while respecting stability is a computable measure of control-based *trust* in the teacher: its establishment and update, its incorporation into controller updates, and its role in shared control during physical interaction.

Our approach first derives an initial control behavior via optimal control, then engages a human teacher to provide physical guidance for corrective demonstration, and finally verifies that the controller produced as a result of the inferred corrections is in fact stable. This verification furthermore is used to estimate a measure of trust in the teacher—relevant for real-world systems learning from multiple teachers with perhaps varying levels of proficiency in providing instruction. Whether to incorporate the teacher's instruction is decided by this trust measure, as well as whether to cede control to the teacher during correction. Given state x, input u, state perturbation z and input perturbation v, we assume that a system has dynamics $\dot{x} = f(x, u)$ where f is piecewise smooth and that $\dot{z} = \frac{\partial f}{\partial x}(t)z + \frac{\partial f}{\partial u}(t)v$ is at least sometimes a locally controllable and observable system.

Operator Demonstration Interpretation Given some initial trajectory $q_i(t)$ (where $x = (q, \dot{q})$) and control input $u_i(t)$ of a system, the operator is allowed to perturb the trajectory by $\eta_i(t)$ to obtain a new trajectory $q_i(t) + \eta_i(t)$. One natural option for how to interpret $\eta_i(t)$ is to infer from it the

objective of the operator; this is called inverse optimal control (IOC) or imputed control (Keshavarz, Wang, and Boyd 2011). Viewing $\eta_i(t)$ as an incremental, local improvement to the original trajectory thus makes the computation feasible and leads to an implementable embedded system. Another option is to treat the operator's perturbation $\eta_i(t)$ as an *exaggeration*. Exaggeration plays a fundamental role in human motion learning. Machine learning techniques like novelty detection and statistical variance can be used to identify exaggeration, and techniques like dimensionality reduction can be used to infer its salient characteristics (i.e. a relevant subset or projection of the state dimensions).

Human Operator Evaluation In the case of physical interaction, trust between the system and operator becomes substantially more important because injury is a possibility. We propose that, in addition to metrics such as improved performance, *a measure of trust could include the (i) stability of a perturbed trajectory and (ii) second-order optimality of perturbations.* The idea is that a human instructor should never propose new trajectories outside the domain of attraction of a stabilizing controller. Moreover, for a subset of canonical motions the locally "best" perturbation can be derived analytically by solving an infinite-dimensional local quadratic model (Hauser and Saccon 2006). The machine thus has a metric, intimately tied to its dynamics, by which it can decide whether or not an operator can be trusted.

Computing Control-Based Trust Consider Figure 1, which illustrates an iteration of an optimal control algorithm. A perturbation $\eta_i(t)$ applied by the operator might need to be scaled down by step size γ if the local feedback law obtained for $q_i(t)$ is to be able to stabilize the perturbation (Armijo 1966). Local stability implies that there exists a γ_k that guarantees the resulting trajectory will be stable, and how large γ is depends on $\eta_i(t)$ (for a Newton step, $\gamma = 1$ will work).

Hence, we can use γ as a computable measure of quality for η_i and accordingly also of trust (τ_a). Moreover, in some scenarios with a known reference trajectory, the optimal perturbation can be computed explicitly and compared to the operator's perturbation, as another factor in the measure of trust (τ_b). Lastly, note that the new, perturbed trajectory is initially stabilized using the previous controller, but then a new set of controllers about the new trajectory are computed. If this new controller is badly conditioned (e.g. near singular) then this can also be incorporated into the estimate of trust (τ_c).

A single metric for trust τ will be established for each human operator, which is learned over time and updates with each trust interaction (i.e. $\{\tau_a^0, \tau_b^0, \tau_c^0, \tau_a^1, ...\}$). Straightforward formulations for τ might consider only the most recent trust interaction (i.e. $\tau^t = \tau_c^t$) or simply compute a running average (i.e. $\tau^t = \tau^{t-1} + \frac{1}{t}((\tau_a^t + \tau_b^t + \tau_c^t) - \tau^{t-1})$). Other formulations might weigh trust interactions temporally—for example to give higher weight ($\alpha \ll \frac{1}{2}$) to past interactions, for a slowly adapting metric (e.g. $\tau^t = \alpha(\tau_a^t + \tau_b^t + \tau_c^t) + (1-\alpha)\tau^{t-1}), \alpha \in [0,1]$). Still others might weight by trust type—for example to give higher weight to τ_b than τ_c, indicating that it is worse to demonstrate outside of the basin of attraction than to demonstrate a badly conditioned controller (e.g. $\tau^t = \alpha(w_0\tau_a^t + w_1\tau_b^t + w_2\tau_c^t) + (1-\alpha)\tau^{t-1}), \alpha \in [0,1], \sum_i w_i = 1$). A worse-case choice also could be made, using $\tau^t = \min\{\tau_a^t, \tau_b^t, \tau_c^t\}$.

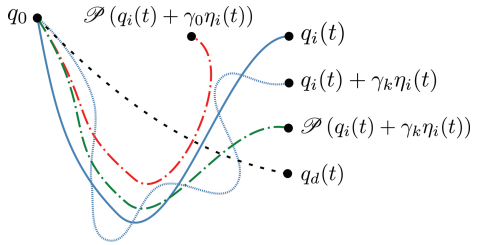

Figure 1: Starting with a feasible trajectory $q_i(t)$ at iteration i, and desired trajectory $q_d(t)$, a descent direction $\eta_i(t)$ can be computed (or demonstrated). From the projection $\mathscr{P}$ a feasible trajectory can be computed via $\mathscr{P}(q_i(t) + \gamma\eta_i(t))$ (where γ is a step size), which may be outside the basin of attraction for the projection if γ is too large (red line). However, it is guaranteed that there exists a γ_k small enough so that the resulting curve $q_i(t) + \gamma_k\eta_i(t)$ is within the basin. The largest such γ_k is a good measure of the quality of $\eta_i(t)$.

Control Authority Transfer Once a measure of trust has been established, nontrivial instruction is allowed to commence. The system only issues a warning if the other states are destabilizing, and the threshold on issuing this warning goes up with trust. After the motion is complete, the perturbed trajectory ξ is stored and $\mathscr{P}(\xi + \gamma\eta)$ can be computed (Fig. 1), looking for the largest γ that is still stable. Note that, with regards to *exaggeration*, this means the operator *can* exaggerate instruction and the choice of γ will scale the result back down, but the system will only *allow* exaggeration after significant trust has been established.

This paper has proposed a computable notion of trust that allows an embedded system to assess the safety of instruction, as a step towards addressing the question: How much should a person be allowed to interact with a robot?

Acknowledgements This paper is based on work supported by the National Science Foundation under award CNS-1329891 CPS: Synergy: Collaborative Research: Mutually Stabilized Correction in Physical Demonstration. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Argall, B. D.; Chernova, S.; Veloso, M. M.; and Browning, B. 2009. A survey of robot learning from demonstration. *Robotics and Autonomous Systems* 57(5):469–483.

Armijo, L. 1966. Minimization of functions having lipschitz continuous first-partial derivatives. *Pacific Journal of Mathematics*.

Hauser, J., and Saccon, A. 2006. A barrier function method for the optimization of trajectory functionals with constraints. *IEEE Int. Conf. on Decision and Control (CDC)* 864–869.

Keshavarz, A.; Wang, Y.; and Boyd, S. 2011. Imputing a convex objective function. In *Proc. of IEEE Multi-Conf. on Systems and Control*, 613–619.