Task-Driven Control via Information Bottlenecks

Vincent Pacelli

State-of-the-art techniques for controlling robotic systems typically rely on accurately estimating the full state of the system and maintaining rich geometric representations of their environment. For example, a common approach to navigation is to build a dense occupancy map produced by scanning the environment and to use this map for planning and control. Similarly, control for walking or running robots typically involves estimating the full state of the robot (e.g., joint angles, velocities, and position relative to a goal location). However, such representations are often overly detailed when compared to a *task-driven representation.*

One example of a task-driven representation is the "gaze heuristic" from cognitive psychology [1–3]. When attempting to catch a ball, an agent can estimate the ball's position and velocity, model how it will evolve in conjunction with environmental factors like wind, integrate the pertinent differential equations, and plan a trajectory in order to arrive at the ball's final location. In contrast, cognitive psychology studies have shown that humans use a dramatically simpler strategy that entails maintaining the angle that the human's gaze makes with the ball at a constant value. This method reduces a number of hard-to-monitor variables (e.g., wind speed and air resistance) into a single easily-estimated variable. Modulating this variable alone puts the human in position to catch the ball.

The gaze heuristic example highlights two primary advantages of using a task-driven representation. First, a control policy that uses such a representation is more efficient to employ online since fewer variables need to be estimated. Second, since fewer variables need to be estimated to implement the control strategy, there are fewer sources of measurement uncertainty. While one can sometimes manually design task-driven representations for a given task, we currently lack a principled theoretical and algorithmic framework for *automatically synthesizing* such representations. The goal of this work is to develop such an algorithmic approach.

Related Work

By far, the most common approach for controlling robotic systems with nonlinear dynamics and partially observable states is to *independently* design an estimator for the *full state* (e.g., a Kalman filter [4]) and a controller that assumes perfect state information. This strategy is optimal for Linear-Quadratic-Gaussian (LQG) problems due to the *separation principle* [5]. However, it can produce a brittle system due to errors in the state estimate for nonlinear systems when the separation principle does not hold. Moreover, to employ this strategy, all available sensor information must be fully exploited to produce an accurate state estimate. This is often a computationally intensive process when the system state is high dimensional (e.g. when the state contains a representation of the environment, such as an occupancy map).

Historically, the work on designing *information constrained* controllers has been pursued within the networked control theory literature [6–8]. Recently, the optimal joint design of data-efficient sensors and performant controllers has also been explored beyond network applications. Some recent approaches include using control policy information content as a surrogate for input costs [9–13] and attempting to select minimum subsets of sensors to provide a requested amount of state information [14–16]. However, the most pertinent body of work is recent attempts to design control policies that minimize the mutual information between system states and control inputs subject to some cost constraints [17–20]. These works have a similar goal as the current work — constructing performant, low-information control policies — but do not consider the robustness and efficiency benefits of such a policy.

Present and Future Research Overview

Preliminary results due to this line of inquiry are under review for publication in the proceedings of the 2019 IEEE International Conference on Robotics and Automation (ICRA) and are available on ArXiv [21]. The main technical contribution of this paper is to formulate the synthesis of task-driven representations as an optimization problem using *information bottleneck theory* [22]. In summary, the stochastic control $\pi_t(u_t|x_t)$ is decomposed into two stochastic maps: $q_t(\tilde{x}_t|x_t), \pi_t(u_t|\tilde{x}_t)$. The following optimization problem is then solved

$$\underset{\substack{q_t(\tilde{x}_t|x_t),\\ \pi_t(u_t|\tilde{x}_t)}}{\text{minimize}}\ I(x_t;\tilde{x}_t) \qquad \text{subject to} \qquad \mathbb{E}\left(\sum_{t=1}^{T-1}(c_t(x_t,u_t)) + c_T(x_T)\right) \leq c_{max}$$

where $c_1, \ldots, c_T$ are cost functions and $I(x_t;\tilde{x}_t)$ is the mutual information. The variables $\tilde{x}_t$ are *task-relevant* in that they contain the minimal amount of information necessary to perform control at the desired fidelity. Effectively, they create a bottleneck between the state and control input that filters out aspects of the state that are not relevant for control. Moreover, these variables can be directly estimated online instead of the higher-dimensional full system state. This approach is inspired by the aforementioned [17–20], but goes further to include characterizations of the online robustness to *unmodeled* measurement noise and computational benefits. This analysis was done both formally by connecting the studied optimization problem to the theory of *risk minimization* and experimentally through numerical system of a variety of control problems. This includes control of the nonlinear spring-loaded inverted pendulum (SLIP) model shown in Figure 1.

Looking forward, the results presented in this paper prompt the investigation of a number of open questions. First, empirically we found that c_{max} describes the tradeoff between a completely closed-loop and completely open-loop policy. A natural next step is to analyze this tradeoff theoretically and then connect it formally to the cost variance or a similar metric of robustness. Next, an issue with the described formulation is that the task-relevant variables $\tilde{x}_t$ are not Markovian, specifically due to the fact that $q_t(\tilde{x}_{t+1}|\tilde{x}_t, u_t, x_t) \neq q_t(\tilde{x}_{t+1}|\tilde{x}_t, u_t)$. Designing the task-relevant variables to be Markovian allows for more accurate online estimation of the task-relevant variable without relying on higher-dimensional state data. Another direction for this work is to incorporate active sensing into the design of the task-relevant variables and control policy. In this setting, the control policy steers the agent to low-cost states that allow for very accurate estimation of the task-relevant variables allows the agent to actively exploit the most salient features of its environment to complete its task in a robust and efficient manner.

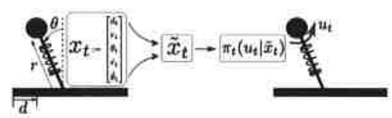

Figure 1: A schematic of the technical approach in [21]. We seek to synthesize (offline) a minimalistic set of task-relevant variables (TRVs) $\tilde{x}_t$ that create a bottleneck between the full state x_t and the control input u_t. These TRVs are estimated online in order to apply the policy π_t. We demonstrate our approach on a SLIP model whose goal is to run to a target location. Our approach automatically synthesizes a *one-dimensional* TRV $\tilde{x}_t$ sufficient for achieving this task.

References

[1] G. Gigerenzer, *Gut feelings: The intelligence of the unconscious.* Penguin, 2007.

[2] P. McLeod, N. Reed, and Z. Dienes, "Psychophysics: How fielders arrive in time to catch the ball," *Nature*, vol. 426, no. 6964, p. 244, 2003.

[3] D. M. Shaffer, S. M. Krauchunas, M. Eddy, and M. K. McBeath, "How dogs navigate to catch frisbees," *Psychological Science*, vol. 15, no. 7, pp. 437–441, 2004.

[4] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics.* MIT press, 2005.

[5] B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods.* Courier Corporation, 2007.

[6] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[7] A. S. Matveev and A. V. Savkin, *Estimation and control over communication networks.* Springer Science & Business Media, 2009.

[8] T. Tanaka, K.-K. K. Kim, P. A. Parrilo, and S. K. Mitter, "Semidefinite programming approach to Gaussian sequential rate-distortion trade-offs," *IEEE Trans. Automat. Control*, vol. 62, no. 4, pp. 1896–1910, 2017.

[9] N. Tishby and D. Polani, "Information theory of decisions and actions," in *Perception-Action Cycle.* Springer, 2011, pp. 601–636.

[10] J. Grau-Moya, F. Leibfried, T. Genewein, and D. A. Braun, "Planning with information-processing constraints and model uncertainty in Markov decision processes," in *Joint European Conf. Machine Learning and Knowledge Discovery in Databases.* Springer, 2016, pp. 475–491.

[11] E. Todorov, "Efficient computation of optimal actions," *Proc. Nat. Academy of Sciences*, vol. 106, no. 28, pp. 11 478–11 483, 2009.

[12] D. A. Braun, P. A. Ortega, E. Theodorou, and S. Schaal, "Path integral control and bounded rationality," in *IEEE Symposium on Adaptive Dynamic Programming And Reinforcement Learning (ADPRL).* IEEE, 2011, pp. 202–209.

[13] G. Williams, N. Wagener, B. Goldfain, P. Drews, J. M. Rehg, B. Boots, and E. A. Theodorou, "Information theoretic MPC for model-based reinforcement learning," in *IEEE Int. Conf. Robotics and Automation*, 2017.

[14] J. L. Williams, J. W. Fisher III, and A. S. Willsky, "Performance guarantees for information theoretic active inference," in *Artificial Intelligence and Statistics*, 2007, pp. 620–627.

[15] V. Tzoumas, L. Carlone, G. J. Pappas, and A. Jadbabaie, "Control and sensing co-design," *arXiv preprint arXiv:1802.08376*, 2018.

[16] L. Carlone and S. Karaman, "Attention and anticipation in fast visual-inertial navigation," in *IEEE Int. Conf. Robotics and Automation.* IEEE, 2017, pp. 3886–3893.

[17] A. Achille and S. Soatto, "A separation principle for control in the age of deep learning," *Annu. Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, p. null, 2018. [Online]. Available: https://doi.org/10.1146/annurev-control-060117-105140

[18] R. Fox and N. Tishby, "Minimum-information LQG control part i: Memoryless controllers," in *55th Conf. Decision and Control.* IEEE, 2016, pp. 5610–5616.

[19] ——, "Minimum-information LQG control part ii: Retentive controllers," in *Proc. 55th Conf. Decision and Control.* IEEE, 2016, pp. 5603–5609.

[20] T. Tanaka, P. M. Esfahani, and S. K. Mitter, "LQG control with minimum directed information: Semidefinite programming approach," *IEEE Trans. Automat. Control*, vol. 63, no. 1, pp. 37–52, 2018.

[21] V. Pacelli and A. Majumdar, "Task-driven estimation and control via information bottlenecks," *Extended version*, 2018. [Online]. Available: https://irom-lab.princeton.edu/publications/

[22] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *Proc. 37th Annu. Allerton Conf. Communication, Control, and Computing*, 1999.