Bridging Inference, Control and Driving: ICDNS

MSci/MSc

Nick Jones

nick.jones@imperial.ac.uk

The course so far

We looked at how biological systems might perform inference and connected to samplers.

We looked at instantiations of deterministic control in two different chemical circuits.

We moved to studying optimal control and closed by looking at stochastic optimal control and at tractable versions.

We will move to considering ways in which biological systems power themselves and how to analyse this.

First we will connect inference, control and driving. Support reading for this lecture material can be found in Refs. [1, 2] (with more context in Refs. [3, 4, 5]). This bridge is a lively area with contributions by authors including Kappen, Toussaint, Todorov and their collaborators.

To cover

To come

1. briefly introduce Markov Decision Problems.
2. consider a Kullback-Leibler divergence-based class of costs for MDPs which are linearly solvable.
3. consider a particular inference task and show that its solution is the same as calculating the preceding optimal cost-to-go.
4. show how there is a minimum energy cost for transforming one distribution into another and this is the same cost as above.
5. connect this back to biological systems

We will thus find that a particular cost-function is the minimum energy cost of control and that this minimum energy cost has a close connection to a natural inference task: bridging inference, control and energy supply. This will build on remarks I made at the close of the last lecture (but will consider a slightly simpler discrete time setting).

A note on dual control

Thus far we have considered optimal control when parameters and dynamics are known and the state is known or unknown. In principle there could be unknown parameters. One way of thinking about this is in terms of priors over unknown parameters. If these priors are parameterized then learning more about unknowns will change the values of the parameters. We can thus think about control as steering not just the state x but also parameters governing our beliefs about unknowns θ. The optimal control will thus yield a sequence of (vector) pairs ...$(x_t, \theta_t), (x_{t+1}, \theta_{t+1})$.... It might be globally optimal to start off (when we have uncertain beliefs) by implementing controls that lead to refined beliefs about the relevant parameters θ but which might not help change our state x towards a desired location (an exploration before exploitation approach) [5].

A note on dual control II

It might be globally optimal to start off (when we have uncertain beliefs) by implementing controls that lead to refined beliefs about the relevant parameters θ but which might not help change our state x towards a desired location (an exploration before exploitation approach). E.g. you spend time and energy finding out how the gearbox works before going on the motorway (or babies don't have holding forks as a priority). In this dual control we expect our control moves to help us both learn/infer and obtain our objective in x space. In nature x might encode the orientation of a bacterium and its velocity but θ is *also physical* and might be encoded by the degree of phosphorylation of some of its proteins. When we analyse a biological system *the distinction between x and θ might thus seem arbitrary.*

Orientation

There is a connection between Kalman filters and linear quadratic Gaussian control which can be explored here: [1].

Introduce a natural class of control problems.

Introduce a particular inference problem.

Show that the inferred probability satisfies the same equation as (an exponentiated version of) the optimal cost-to-go.

Markov Decision Problems

An MDP is an optimal control problem with finite controls and states and probabilistic state transitions (which are conditional on the controls).

We can suppose that our uncontrolled system has dynamics: $p(x'|x)$.

We suppose that the controlled dynamics yields a different distribution $h(x'|x, u(x))$ we will refer to this *conditional distribution* as $u(x'|x)$ for short (tracking the notation in [1] as we mostly will for the discussion of the duality Todorov pointed out).

A natural choice of cost function

We will consider costs of the form:
$w(x, u, t) = q(x) + D(u(x'|x)||p(x'|x))$ (and will be considering final costs $W(x, T) = 0$).
$D(u(x'|x)||p(x'|x))$ is the Kullback-Leibler Divergence. In the discrete case this is $\sum_{x'} u(x'|x) \log \frac{u(x'|x)}{p(x'|x)}$. The form of $q(x)$ is unconstrained.
We will relate this, at the close of the (extended) lecture, to a lower bound on the amount of work that needs to be done to go from $p(x'|x)$ to $u(x'|x)$. The Kullback-Leibler Divergence is not symmetric and has a diversity of elegant interpretations.
We now discuss the case $u(x'|x) = p(x'|x)$ and the case x'' where $u(x''|x) \neq 0$ but $p(x''|x) = 0$.

Linearly Solvable MDPs

Optimal cost-to-go is thus:

$$J(x,t) = \min_{u(x'|x)} \left[q(x) + D(u(x'|x)||p(x'|x)) + \mathbb{E}_{u(x'|x)}[J(x',t+1)] \right]$$

where $\mathbb{E}_{u(x'|x)}$ encodes an expectation when x' is distributed as $u(x'|x)$.

Linearly Solvable MDPs II

$$J(x,t) = \min_{u(x'|x)} \left[q(x) + D(u(x'|x)||p(x'|x)) + \mathbb{E}_{u(x'|x)}[J(x',t+1)] \right]$$

Defining $\phi = \mathbb{E}_{p(x'|x)} \exp(-J(x',t+1))$ then a small amount of rearrangement of the above yields:

$$J(x,t) = q(x) - \log\phi + \min_{u(x'|x)} \left[D\left(u(x'|x)|| \frac{p(x'|x)\exp(-J(x',t+1))}{\phi} \right) \right]$$

We can minimize this by setting $u'(x'|x,t) = \frac{1}{\phi} p(x'|x)\exp(-J(x',t+1))$ which sets the KL-divergence term to zero:

$$J(x,t) = q(x) - \log \mathbb{E}_{p(x'|x)} exp(-J(x',t+1))$$

.

Linearly Solvable MDPs III

$u'(x'|x,t) = \frac{1}{\phi} p(x'|x)\exp(-J(x',t+1))$ has a simple interpretation as rescaling the transition dynamics, under passive distribution $p(x'|x)$, by how expensive the states x' will prove to be given $J(x',t+1)$.
We can rewrite the optimal cost-to-go as:

$$z(x,t) = \exp(-q(x))\mathbb{E}_{p(x'|x)}[z(x',t+1)]$$

where $z(x,t) = \exp(-J(x,t))$. This is linear in z and much easier to solve than unconstrained MDPs [4].

Orientation

(Tick) Introduce a natural class of control problems.
Introduce a particular inference problem.
Show that the inferred probability satisfies the same equation as (an exponentiated version of) the optimal cost-to-go.

An HMM forecasting task: the chance of a string of free lunches?

A traveller moves from city state to city state according to some transition matrix $p(x'|x)$ (and later we'll associate a transition cost for these moves).
Each city state x has different levels of meanness $q(x)$ such that the chance of her getting a free lunch $p(y_t = 0|x_t) = \exp(-q(x_t))$ is higher or lower depending on the meanness of the city (x_t is the state at time t). The probability of her getting a free lunch will be related to a state cost and a 'free-lunch event' tells us something about the state x.
If all that we can observe is whether her lunch is free or not through time then this is a Hidden Markov Model. We undergo Markov dynamics in x according to $p(x'|x)$ but x is hidden and we only observe the binary variable y with probability of $y = 0$ being $\exp(-q(x_t))$.

An HMM forecasting task: the chance of a string of free lunches? II

A traveller moves from city state to city state according to some transition matrix $p(x'|x)$.
Each city state x has different levels of meanness, $q(x)$, such that the chance of her getting a free lunch is
$p(y_t = 0|x_t) = exp(-q(x_t))$.
HMM forecasting task: Given that she starts off at x, t, and given $q(x)$ and $p(x'|x)$ what is the chance that she has only free lunches until time T? I.e. given the current state predict one particular set of future observations $y_{t:T} = \mathbf{0}$. I.e. $p(y_{t:T} = \mathbf{0}|x_t)$.
The chance of a sequence of free lunches depends on both the local meanness $q(x)$ (state-cost) and also how likely it is that the traveler's dynamics, $p(x'|x)$, take her to generous places. [It's a forecasting problem but HMM's are generically set up for the inference task of, given observed free lunches, try to infer quantities like x, $q(x)$ and $p(x'|x)$.]

Probability of the total free lunch run?

Define $r(x,t) = p(y_{t:T} = \mathbf{0}|x_t)$.
It follows that

$$r(x,t) = p(y_t = 0|x_t) \sum_{x'} p(x'|x_t) r(x', t+1)$$

or

$$r(x,t) = exp(-q(x)) \sum_{x'} p(x'|x) r(x', t+1)$$

or

$$r(x,t) = exp(-q(x)) \mathbb{E}_{p(x'|x)} r(x', t+1)$$

this has the same form as the equation we derived for the optimal cost-to-go:

$$z(x,t) = exp(-q(x)) \mathbb{E}_{p(x'|x)} [z(x', t+1)]$$

where $z(x,t) = exp(-J(x,t))$.

Intuition behind duality

We found that $r(x,t) \propto exp(-J(x,t))$. Meanness $q(x)$ is state cost $q(x)$.
If $r(x,t)$ is large ($J(x,t)$ small): likely to have a run of free lunches = low cost/ easy to control to obtain this cheap outcome. Passive (free) dynamics takes me to places where I'm likely to get lunch for free (low meanness $q(x)$).
If $r(x,t)$ is small ($J(x,t)$ large): Unlikely to have all free lunches = extensive intervention (departures from passive dynamics) required to obtain such low state cost trajectories. Controlled (expensive) dynamics is required for many free lunches.
Initial conditions x which are unlikely, under passive dynamics, to move through a sequence of low-cost configurations are ones which require a high control cost.
In statistical physics we are accustomed to seeing exponentiated energies being related to probabilities; we will refine this.
What if $q(x) = 0$? (equivalent to setting controlled dynamics to be the passive dynamics)

Orientation

(Tick) Introduce a natural class of control problems. MDPs with KL costs. These can naturally be related to the formulation of Kappen and extended to the continuous time setting.
(Tick) Introduce a particular inference problem. The probability of a particular sequence of configurations given costs for each state and certain transition dynamics.
(Tick) Show that the inferred probability satisfies the same equation as (an exponentiated version of) the optimal cost-to-go. This duality holds in the continuous time setting [1].
Now we will investigate why MDPs with KL costs - as well as yielding simple solutions and being sensibly related to inference tasks - are particularly appropriate. This will be tracing the nice treatment in Ref. [2].

Szilárd Engine

Work done in compressing a single molecule of gas in a cylinder from V to $V/2$. If an ideal gas in isothermal limit then $PV = k_B T$ and $W = -\int_V^{V/2} PdV = k_B T \log 2$. Which half of the cylinder contains the gas is a choice. If we know the side with the molecule then we can recover work by inserting a partition and arranging pulleys etc so that the molecule can collide with the partition to do work. After doing work colliding with the partition the molecule is then uniformly distributed in the cylinder. There is thus an interplay between work done and information about where the particle is (its distribution within the cylinder).

Szilárd-Landauer Correspondence

We will go through the proof showing that $k_B T$ times the Kullback-Leibler divergence between distributions u and π is equal to the free-energy difference between these states.
We suppose we have n system configurations and the i^{th} configuration is assigned the energy $E(i)$ and we have a temperature T. We further assume that configuration i occurs with probability u_i under distribution u. Average energy $\langle E \rangle_u$ under u is thus $\sum_{i=1}^{n} u_i E(i)$ and entropy $H(u)$ is $\sum_{i=1}^{n} -u_i \log u_i$.

Definitions

1. Free energy at temperature T: $F_{E,T}(u) = \langle E \rangle_u - k_B T H(u)$
2. Partition function: $Z_{E,T} = \sum_{i=1}^{n} exp(-\frac{E(i)}{k_B T})$
3. Gibbs distribution: $\pi_{i,E,T} = \frac{1}{Z_{E,T}} exp(-\frac{E(i)}{k_B T})$

We will now discuss the interpretation of the free energy.

Szilárd-Landauer Correspondence II

Theorem

Szilárd-Landauer Correspondence
Given the Gibbs distribution π and distribution u then
$F_{E,T}(u) - F_{E,T}(\pi) = k_B T D(u||\pi)$

We can thus interpret taking a system with equilibrium distribution $\pi = p$ and converting it into one distributed as u as requiring energy $\geq k_B T D(u(x'|x)||p(x'|x))$. [Sometimes physical understanding will hand us the set of $E(i)$ but generically we can, given a distribution π, construct a corresponding T and set of $E(i)$.]

Szilárd-Landauer Correspondence Proof

$F_{E,T}(u) = \langle E \rangle_u - k_B TH(u) = \sum_{i=1}^{n}(u_i E(i) + k_B T u_i \log u_i)$. From the Gibbs distribution we can rearrange for E_i:
$E_i = -k_B T \log \pi_i - k_B T \log Z$ and then eliminating E_i from $F_{E,T}(u)$ and rearranging yields:
$F_{E,T}(u) = k_B TD(u||\pi) - k_B T \log Z$.
One can show that $F_{E,T}(\pi) = -k_B T \log Z$ by writing:
$F_{E,T}(\pi) = \langle E \rangle_\pi - k_B TH(\pi) = \sum_{i=1}^{n} \pi_i E(i) + k_B T \pi_i \log \pi_i$ and using $\pi_{i,E,T} = \frac{1}{Z_{E,T}} exp(-\frac{E(i)}{k_b T})$.

Dualities and the Minimal cost of MDP control

Given a controlled distribution $u(x'|x)$ and passive distribution $p(x'|x)$ we know that the minimum energy cost of this control is $\geq k_B T D(u(x'|x)||p(x'|x))$.

Thus the cost of control of *MDPs* is *always* greater than or equal to $k_B T D(u(x'|x)||p(x'|x))$. A cost function with this term is thus always a lower bound on the actual cost.

The choice of cost function $w(x, u, t) = q(x) + D(u(x'|x)||p(x'|x))$ is thus particularly natural and the duality between inference and control particularly relevant (since it occurs for a cost function which is arguably fundamental).

Implications for Natural Systems

We can see that natural systems will have a cost function which will also be bounded from below. We have previously seen a connection between Bayesian inference and samplers. We have discussed how some chemical and neural systems can act like samplers. We have now seen a connection between a set of inference tasks and a well motivated suite of optimal control tasks. Constructing chemical samplers suitable for optimal control is a wide-open research problem, but despite this, we know something about their physical limits.

Bibliography

[1] E. Todorov, General duality between optimal control and estimation. In proceedings of the 47th IEEE Conference on Decision and Control, 4286, 2008. Free online.

[2] M. Gopalkrishnan, The Hot Bit I: The Szilárd-Landauer Correspondence, arXiv: 1311.3533. Free online.

[3] E. Todorov, Optimal control theory, In Bayesian Brain: Probabilistic Approaches to Neural Coding, Doya K at al (eds), chap 12, 269, MIT Press 2006. Free online.

[4] E. Todorov, Efficient computation of optimal actions, PNAS, 11478, 83, 2009. Free online.

[5] H.J. Kappen, Optimal conrol theory and the linear Bellman Equation. Book chapter. Free online.