Controlling a Sepsis Simulation with PILCO, a Model-learning Controller

Thomas Desautels, Brenden Petersen, Chase Cockrell*, Claudio Santiago, Gary An*, and Daniel Faissol

LLNL, *University of Chicago

Aug 8, 2018
Outline

- What is sepsis?
- Cytokine mediation
- Agent-based modeling
- PILCO
- Results
- Conclusions and Future Work

Apply PILCO to control an agent-based model of sepsis
What is sepsis?

- A harder question than it sounds
  - A “life-threatening organ dysfunction due to a dysregulated host response to infection” [1].
  - Traditionally thought of as over-inflammation [2], but this has changed.

- Statistics are difficult to pin down, but:
  - affects roughly 2% of hospital inpatients [4];
  - results in in-hospital mortality of approximately 10% [1, 4]; and
  - costs the US healthcare system approximately $20 billion annually [1].

- Current treatment is largely supportive.
Cytokine Mediation

- Cytokines are key players in the immune response
  - Signaling molecules that provide main interface between cells in immune response

- Can we modulate the cytokines to mitigate sepsis?
  - Previous efforts to treat sepsis via cytokine mediation have failed

- Hypothesis:
  - Cytokine mediation requires a more sophisticated control strategy
Agent-based Modeling

- System is modeled as a collection of “agents” that follow designer-specified rules, yielding emergent behavior
  - Rules can include switching, other behaviors not implementable in DEs

- Simple examples:
  - Particles in a box, following laws of Newtonian Physics → ideal gas model
  - Rabbits and grass
    - Rabbits move to seek grass, eat grass, seek other rabbits, reproduce
    - Grass replicates itself
    - → 2-state population dynamics

```python
class Agent():
    def step():
        # modify self, environment, other agents
        self.posnew = self.pos + self.v
        if self.posnew through wall:
            self.posnew, self.v = wallreflect(self.pos, self.v)
        self.pos = self.posnew

T = 100; t = 0; agents = [...]; done = False
while not done:
    for agent in agents:
        agent.step()  # could set done = True
    t = t + 1
    if t > T:
        done = True
```

Simple examples:
- Rabbits and grass
  - Rabbits move to seek grass, eat grass, seek other rabbits, reproduce
  - Grass replicates itself
  - → 2-state population dynamics
Sepsis ABM

Immune Cells (Mobile Agents)

PMN (Neutrophil)

Monocyte

TH0

TH1

TH2

Toxins and Cytokines (Environment)

cytotoxin

endotoxin

Infection

“Oxy” (Endothelial Cell)

Spatially Fixed entities

Cell-to-cell messages via cytokines: Nearly bipartite

See [5, 6] for sepsis ABM implementation
Developed by M. Deisenroth, now of Imperial College London, under Carl Rasmussen (Cambridge) [7]

Traditional control theory:
- Use first-principles models of a system’s dynamics
- Control policy: observes system and emits control actions
- Optimize policy with respect to a cost function on the system’s behavior
- Doesn’t work if you don’t have a system model

Key PILCO ideas:
- Learn a data-driven probabilistic model of the system’s transitions $s_t, a_t \rightarrow s_{t+1}$
- Closed-form approximate gradients of expected cost $J$ with respect to the policy parameters $\theta$.
Assume there is learnable structure in the system dynamics

- PILCO: Probabilistic Inference for Learning COntrol

- Policy: $a_t = \pi(s_t | \theta)$

- Sepsis ABM: $s_{t+1} \leftarrow s_t, a_t$

- Apply policy

- Update Gaussian process dynamics model

- Re-optimize policy: $\nabla_{\theta} E_s [J(s, \pi(s | \theta))]$
\[ s_{t+1} = s_t + f(s_t, a_t) + \omega, \quad f \sim \text{GP}(0, k_{SE}(s, s')) , \omega \sim \mathcal{N}(0, \Sigma_n) \]

\[ a_t = \pi(s_t | \theta) = a_{\text{max}} \sigma \left( \sum_{i=1}^{n} \alpha_i \exp \left( -\frac{1}{2} (s_t - m_i)'\Lambda^{-1} (s_t - m_i) \right) \right) \]

\[ \theta = \{ \alpha_1, m_1, ..., \alpha_n, m_n, \Lambda^{-1} \} \]

Update \( \theta \) by gradient descent on the approximate expected cost \( J(\theta) \).

\[ C(s) = 1 - \exp(-1/2 (s - g)'Q(s - g)) \]

\[ J(\theta) = \mathbb{E}_{p(s_{t+1} | s_t, \pi(s_t | \theta))p(s_0)} \left[ \sum_{t=1}^{T} C(s_t) \right] \]

- Passing uncertain \( p(s_t | s_{t-1}, a_{t-1}) \) through GP \( f \) results in non-Gaussian \( p(s_{t+1} | s_t, a_t) \); repeated Gaussian approximations by moment matching.
Experiment 1
PILCO applied to a simplified sepsis model

- Fully deterministic 5-state, nonlinear dynamical system with one control
  - Hand tuned:
    - If no infection and small health damage, slowly recovers
    - If any infection, it begins to grow
    - Small region of stability around the goal state (full health, zero infection; lower right) results: ~45 of 156 selected initial conditions “saved”

- Results:
  - Linear optimal identification and control “save” 85
  - PILCO-designed controllers “save” up to 147
Experiment 2
PILCO applied to the sepsis ABM

- Stochastic, 101x101 grid 6 cell types (agents), 12+ cytokines and diffusible chemicals; 12 controls.

- For PILCO:
  - $s_t$ represented as a spatial average of health state, infection, 3 diffusible molecules, one cell count.
  - Controls: one pro-inflammatory cytokine, and one anti-inflammatory
  - PILCO finds a policy that saves 2471/2500 randomly selected patients from a cohort with approximately 50% mortality in the absence of control.
Experiment 2
PILCO applied to the sepsis ABM

PILCO finds a simple *pro-inflammatory* policy that is effective for many virtual patients
Historically, sepsis has been viewed as a hyper-inflammatory disorder, and so decreasing inflammatory cytokines and increasing anti-inflammatory cytokines seems logical.

The highly-effective policy found with PILCO is pro-inflammatory.

Finding counter-intuitive policies is the designed intention of the project and a major advantage of simulation.

Is the model “wrong?”
- State of knowledge embodied by the ABM not valid for the states encountered in simulation?
- Also applying PILCO and DRL to other versions of the model
- The other wing of this project is trying to calibrate the model to experimental data

Are the model and policy “right” and it just hasn’t been tried?
- Testable in animal models
References


