Computing Treatments for Type-1 Diabetes Mellitus

Meysam Bastani\textsuperscript{2}, Kit Chen\textsuperscript{2}, Dr. Michael Bowling\textsuperscript{2}, Dr. Russ Greiner\textsuperscript{2}, Dr. Edmond Ryan\textsuperscript{1}

\textsuperscript{1}Alberta Diabetes Institute \textsuperscript{2}Dept of Computing Science, AICML

\textbf{Introduction}

Patients with type-1 diabetes need their insulin treatment regimes tailored for their individual needs (e.g., age, diet, exercise, stress). Diabeticians have to constantly re-adjust and personalize treatment for Type-1 diabetes as patient’s change.

We hypothesize that computer-based recommendation systems can help doctors serve more patients with more accuracy and dynamic personalization of treatment policies.

\textbf{What can Computers do for Diabetes treatment?}

\textbf{Diabetes is a sequential decision making problem.}

Sequential decision making decisions made now can have both immediate and long-term effects.

Computers can explore and evaluate many possible courses of action--Called policies or treatment policies.

\textbf{Background}

Evaluating the “goodness” of a treatment regime is done with a reward function that maps health outcomes to a number (e.g., QoL, or AIC).

\textbf{Diabetologist Policy}

The diabetologist policy computes the dosage of bolus injection when given the patient’s current blood glucose, and carbohydrate intake

\textbf{Our Task}

- Given: Minute-by-minute blood glucose meter samples as provided by the T1DM Simulator
- Basal (long acting) insulin is set once and simulated as a low dose of basal insulin given once per minute throughout the day.
- Breakfast, Lunch, and Dinner meal sizes and meal times are distributed according to a truncated Normal
- Find:
  - Dosage of bolus (short acting) insulin to take after breakfast, lunch, and dinner.
  - Equivalently, find a policy (the CR and CF parameters) that best regulates blood glucose (that maximizes mean reward in a 3 month time span)

\textbf{T1DM Simulator}

- Simulates the metabolic systems of 10 adult and 10 adolescent type-1 diabetic patients.
- FDA approved to substitute animal trials in pre-clinical testing.
- Each patient has a ground truth optimal treatment policy

\textbf{Why is the task hard?}

- The number of possible actions (treatments) if we have room for 8 bolus injections over the course of a day
  - Each bolus injection involves:
    - A carbohydrate intake
    - A correction factor
    - A basal insulin rate

- The dynamics are complex.

\textbf{Evaluating a policy}

- For each patient, consider 1600 policies with different [CF, CR] parameter settings.
- Compute the mean reward on a single patient (or group of patients) for 9 days and
- Create the following contour plot performance surface.

\textbf{Actor-Critic}

- A reinforcement learning algorithm where the agent is divided into
  - an actor who selects a policy and
  - a critic who evaluates the “goodness” of the chosen action in the current state.
- The separation allows for more efficient action selection (in particular with continuous actions)

\textbf{T1DC}

- Not a reinforcement learning algorithm. Does not adapt policy to new data from patient.
- It is a supervised learning algorithm, which learns a function from labeled training data.
- At training time, the T1DC knows the optimal policy for each training subject.
- Can improve performance very quickly -- in 36 days!

\textbf{SARSA}

- SARSA is a reinforcement learning algorithm
- It is an iterative algorithm that updates the policy based on actions taken.
- Designed for discrete insulin dosages

\textbf{Future Work}

Moving beyond in-silico patients to real patients involves:

- A more realistic model of what happens when a patient becomes hyperglycemic.
- Handling the effects of stress and exercise on a patient’s blood glucose.
- Determining at what point can we be confident that results on our in-silico set reflect that of a real set of patients?