UNIVERSITY OF ALBERTA
ALBERTA DIABETES INSTITUTE

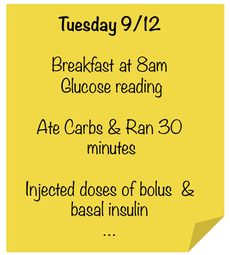
Computing Treatments for Type-1 Diabetes Mellitus

Meysam Bastani[2], Kit Chen[2], Dr. Michael Bowling[2], Dr. Russ Greiner[2], Dr. Edmond Ryan[1]
[1]Alberta Diabetes Institute [2]Dept of Computing Science, AICML

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.

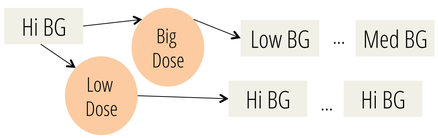
We hypothesize that computer-based recommendation systems can help doctors serve more patients with more accuracy and dynamic personalization of treatment policies.



What can Computers do for Diabetes treatment?

Diabetes is a sequential decision making problem.

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



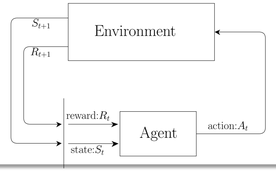
Computers can explore and evaluate many possible courses of action--
Called **policies** or **treatment policies**.

Background

Evaluating the "goodness" of a treatment regime is done with a **reward function** that maps health outcomes to a number (e.g. QaLY, or A1C). For diabetes, one such reward function is **mean reward** over a period of time.

$$\tilde{R}_{t+1} = \frac{\int_{\tau_t}^{\tau_{t+1}} \text{score}(\text{bloodglucose}_\tau)\, d\tau}{\tau_{t+1} - \tau_t}$$

Reinforcement Learning -- a computational approach to goal-directed learning from interaction.



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

Patient-Type	Patient-id(PID)	Body-Weight(Kg)	Age	Optimal-basal	Optimal-Sensitivity(CF)	Optimal-CHO(CR)
Adolescent	1	68.7	18	0.84	0.5	12
Adolescent	2	51	19	0.92	0.5	15
Adolescent	3	44.8	15	0.65	0.5	15
Adolescent	4	49.6	17	0.87	0.5	15
Adolescent	5	47	16	0.72	0.5	15
Adolescent	6	45.5	14	0.88	0.5	15
Adolescent	7	37.9	16	0.78	0.4	17
Adolescent	8	41.2	14	0.56	0.4	17
Adolescent	9	43.9	19	0.61	0.4	17
Adolescent	10	47.4	17	0.79	0.4	17
Adult	11	102.3	61	1.27	0.7	10
Adult	12	111.1	65	1.37	0.7	10
Adult	13	81.6	27	1.43	0.6	13
Adult	14	63	66	0.89	0.5	14
Adult	15	94	52	1.18	0.6	10
Adult	16	66	26	1.72	0.5	12
Adult	17	91.2	35	1.37	0.6	10
Adult	18	102.8	48	1.14	0.7	10
Adult	19	74.6	68	1.13	0.5	13

Diabetologist Policy

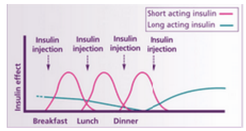
The diabetologist policy computes the dosage of bolus injection when given the patient's current blood glucose, and carbohydrate intake

$$\text{Bolus Injection} = \frac{\text{Current Blood Glucose} - \text{Target Blood Glucose}}{\text{Correction Factor}} + \frac{\text{Carb Intake}}{\text{Carb Ratio}}$$

- **Correction Factor (CF)** and **Carb Ratio (CR)** are patient-specific parameters.

Our Task

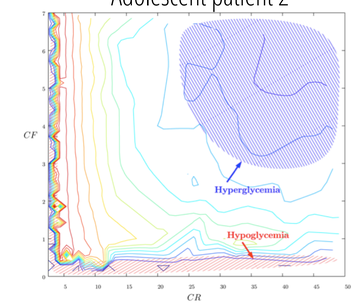
- **Given:**
  - Minute-by-minute blood glucose meter samples as provided by the T1DM Simulator
  - Basal (long acting) insulin is set once and simulated as a low dose of basal insulin given once per minute throughout the day.
  - Breakfast, Lunch, and Dinner meal sizes and meal times are distributed 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)

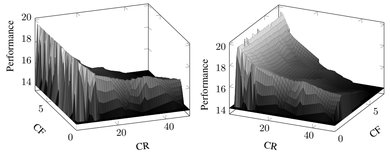


Evaluating a policy

- 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**.

Adolescent patient 2





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!

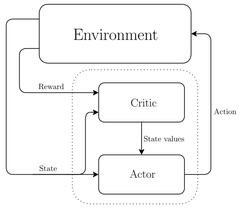
SARSA

- SARSA is a reinforcement learning algorithm
- It is an iterative algorithm that updates the policy based on actions taken.

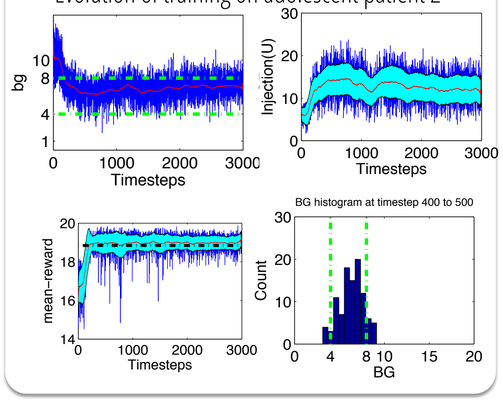
✗ Designed for discrete insulin dosages

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)



Evolution of training on adolescent patient 2



Conclusion

Actor-Critic methods produced the most promising results. Initial policies are reasonable. They find near-optimal policy parameters using just 6 months of simulation time.

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?

Alberta Diabetes INSTITUTE

ALBERTA INNOVATES CENTRE FOR MACHINE LEARNING

Intelligent Diabetes Management Group
http://idm.ualberta.ca/