A MODELING, HEALTH IT, ANALYTICS FRAMEWORK FOR HEALTHCARE MANAGEMENT AND APPLICATION TO DIABETES II

Iakovos M. Katsipis
University of Maryland

John S. Baras
University of Maryland
baras@umd.edu

Keywords: systems engineering**,** decision making**,** trade-offs, disease models**,** care quality

Health care systems in general are complex systems and even systems of systems. **A modern, systems engineering, approach to the modeling, design, construction, operation and maintenance of such systems is lacking [1].** We develop a framework using modern **Model-Based System Engineering** (**MBSE**) methodologies, frameworks and tools. We selected as a focus application **Diabetes Mellitus** (or **Diabetes 2**) for its high impact, as it affects tens of millions of people world-wide. The framework developed is scalable and expandable, can "learn", is linkable to medical databases, can be used by patients, medical personnel, healthcare managers and insurers, provides powerful capabilities for analytics. We focus on a key component of such a system, which is the **reasoning engine** to perform efficiently the required difficult tradeoffs in many key decisions.

A key component is the modeling of disease evolution incorporating medical tests and interventions. We have developed a model for diabetes progression as a Controlled Hidden Markov Chain (CHMC), with three states. ***State 1*** represents the **Healthy** condition, ***State 2*** the **Pre-diabetic** condition, ***State 3*** the **Diabetic** condition of a patient. More complex models can be developed based on more detailed dynamics and progression of the disease supported by clinical data and biochemical models. As the patient goes on with her/his life, various tests are performed periodically to determine the patient's state, and various interventions (treatments) are followed.

We included the following three diagnostic tests: A1C, FPG (fasting plasma glucose), OGTT (oral glucose tolerance test). We included operational characteristics of these tests, as reported in the literature and medical standards; i.e. statistical errors, probabilities of missed detections and false alarms, accuracy, fallacy. We used them in sequential state estimation schemes we developed in our framework. We considered four types of interventions (technology, diet, exercise, medication), and ten specific interventions from these types. A strategy involves a sequence of interventions applied at various times; inspired by similar usage in OR and Engineering. We use a time horizon for the study denoted by *T*; in most of our work *T* is ten years.

We have included several metrics in our studies; our MBSE framework is particularly suitable for multiple metrics. One performance metric that we are interested in, is **cost**. There is a cost for every test and every intervention, denoted by $C_u(u_k) = c_{u_k}$ and $C_\mu(\mu_l) = c_{\mu_l}$. The cost is additive:

$$C_\mu^{total}(i,m_i) = \sum_{t=1}^{N_{T,\Delta}} C_u(u(t)),\ C_\mu^{total}(i,m_i) = \sum_{t=1}^{N_{T,\Delta}} C_\mu(\mu(t)),\ \text{and}\ C^{total}(i,m_i) = C_\mu^{total}(i,m_i) + C_u^{total}(i,m_i)$$

In these sums, the tests and interventions used at each time step of a time history are considered.

We consider **three types of patients**: **"Risk Averse"** patient, **"Risk Indifferent"** patient and **"Risk Taker" patient.** We introduce weights representing the value (or significance) each patient places for being in each state of the model (recall Healthy, Pre-diabetic, Diabetic) V_1^i, V_2^i, V_3^i. These

weights take real nonnegative values between 0 and 1, and they sum to 1. The number of periods that each patient, in each time history, is in each state, is an important health metric. As state transitions depend explicitly on the tests and interventions applied, the three counting statistics below, for each patient time history, constitute a practical health care quality metric.

$O_1^i(m_i)$ = number of periods, patient *i* is at state 1 (Healthy); similarly $O_2^i(m_i), O_3^i(m_i)$ for states 2, 3

Using the weights V_1^i, V_2^i, V_3^i and these counting statistics O_1^i, O_2^i, O_3^i, we can define several metrics for health care quality. For example we consider the following ***Health Care Quality metric***:

$$J_{hc}(i, m_i) = V_1^i * O_1^i(m_i) + V_2^i * O_2^i(m_i) + V_3^i * O_3^i(m_i)$$

Using the model and these metrics, we develop three methods for evaluating healthcare associated with each patient and each time history. The main focus and innovation in these evaluations are systematic tradeoffs (Pareto points). The first method uses the model in an exhaustive straightforward generation of all possible sample paths (time histories) for any number of patients, which we call Evaluation by Monte Carlo simulation (EMCS). The other two methods employ a multi-criteria optimization approach to directly compute the Pareto points and associated selection of tests and interventions. The second method computes the tests and interventions to be applied at time *t* as functions of *x*(*t*), and uses explicitly the state of the disease. It is thus called **Fully Observable**. In the third method the disease state is not available, and instead estimates of the state based on the scores and results of diagnostic tests are recursively computed. This third method is called **Partially Observable**. We show that the second method, called **Fully Observable Multi-criteria Optimization (FOMCO)** saves tremendously in computational time in achieving similar tradeoff analysis as the EMCS Method. Similar computational reduction is shown for the third method, called **Parially Observable Multi-criteria Optimization (POMCO)** These computational savings are important as we apply our methods and tools to very large real-life problems.

We are interested in analyzing the tradeoff between different pairs of sequences $(\boldsymbol{\mu}, \boldsymbol{u})$ from the perspective of total Cost $C^{total}(i, \boldsymbol{\mu}, \boldsymbol{u})$ and average Health Care Quality $\bar{J}_{hc}(i, \boldsymbol{\mu}, \boldsymbol{u})$ metric. The "best" tradeoff points are the **Pareto points**. To compute them using the FOMCO or POMCO methods, we combined the two metrics in a convex combination, resulting in a single criterion stochastic optimization problem. We used Dynamic Programming analytics for both methods.

A key output form our MBSE system, are the Pareto points that describe succinctly the relative value of treatments and tests vs the overall health care quality of a patients time history. Running ECMC with two metrics (and 2-D graphs) for 10,000 patients and 32 runs, took for the whole experiment 783sec. Running ECMS with three metrics (and 3-D graphs) for 100,000 patients took for the whole experiment 1,385 sec. FOMCO outputs directly the Pareto-points and other related information and is very fast. For the same problems that the First Method (EMCS) took 783 sec (two metrics, 10,000 patients, 32 runs) and 1,385 sec (3 metrics, 100,000 patients, 9 runs), our Second Method (FOMCO-SN) took only 2.36 sec on the same laptop. POMCO took only twice the time of FOMCO. The Pareto points and frontiers computed are very similar to those computed with EMCS but at two or more orders faster time!

The system developed has powerful capabilities and can provide powerful healthcare analytics efficiently; primarily enabled by our MBSE approach. We demonstrate these capabilities by posing and answering some interesting realistic questions from the perspective of health care management.

References

1. C. Cassel et al, "Better Health care and Lower Costs: Accelerating Improvement Through Systems Engineering," *Report to the President, Executive Office of the President, President's Council of Advisors on Science and Technology,* May 2014.