Causality and Belief Change

Kinzang Chhogyal[+], Abhaya Nayak[++], Rolf Schwitter[++]
[+] IIS, Griffith University [++] Macquarie University

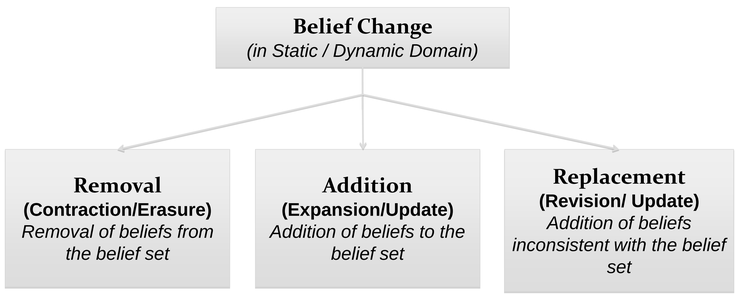
Abstract

Interaction with the world by an agent can lead to the formation of new beliefs, or perhaps, to the confirmation or rejection of existing beliefs. This interaction is also the basis of the formation of rudimentary cause and effect relationships (causal model /structure), and their subsequent improvement. These causal models serve as a tool to evaluate our understanding of the world. We restrict the discussion to one such model in the health domain.

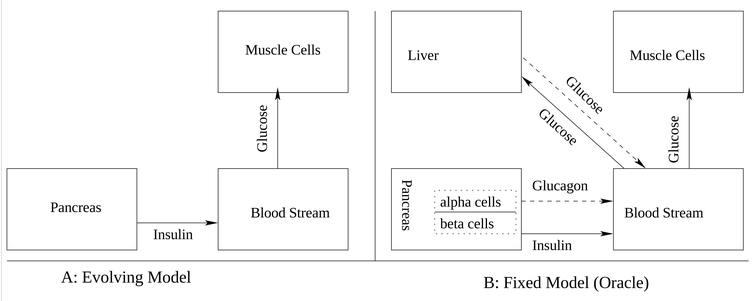
We model causality by exploiting techniques that have been developed in the field of belief replacement (revision and update). Previous attempts have mostly focussed on probabilistic (Bayesian) methods. Instead, we use distance measures as the mathematical foundation in this work.

We start with a very simple and well-understood example from the medical domain. We investigate how an agent that has initially incomplete and/or incorrect (relevant) knowledge can iteratively develop a simple causal model by interacting with an oracle (that represents the "real world"). Given an action, a cycle in this iteration consists of (1) the agent making a prediction, (2) comparison of this prediction with the actual output of the oracle, and (3) subsequent modification in the agent's model. This process is repeated until the model stabilizes.

Belief Change



Knowledge and Real World



Black Box / White Box

Patient Status	Blood Sugar Level		
	Low	*Normal*	*High*
Alert	S1	S2	S3
Not Alert	S4	S5	S6

Actions
Administer Glucose
Administer Insulin

Table 1*. The states of the system that we intend to model and the available actions.*

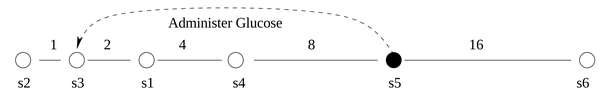


Figure 1*. The black box represents the system we intend to model. The distance (numbers) shown are the actual distance between states.*

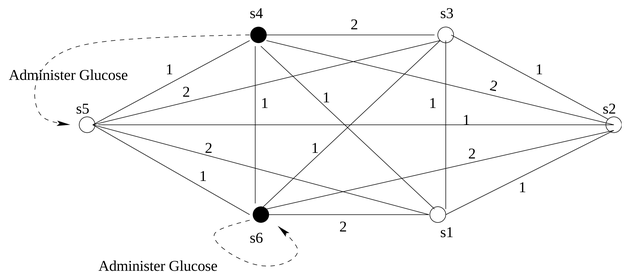


Figure 2. *The white box represents the agent's model of the system. The distance shown are the Dalal distance between states.*

Dynamics

Black Box

Initial State : S5

Administer Glucose

Next State : S3
(patient alert)
Distance to S3 is smaller than to S6

White Box

Initial State : S4, S6

Administer Glucose

Update
Distance from individual states

Next State : S5, S6
(patient not alert)
S5 closest to S4, S6 closest to S6

Revision
Observation and prediction don't match.

Next State: S2, S3
(patient alert)
S2,S3 are closest worlds

Implementation

Figure 3*. Implementation of the white box*

Assumptions

- Finite language (4 atomic sentences)

Distance Measure

- Must be transitive and irreflexive
- We use Dalal Distance: difference in propositional variables

Results & Discussion

- White box stabilizes in avg. 5-6 iterations when actions chosen randomly, transition identical to black box
- Stable model doesn't mean the agent has complete or correct knowledge, maybe in a local minima
- Choice of actions also determine stability
- Distance measure key to finding correct model

Future Work

- Investigation of different distance measures
- Correction of the distance measure itself
- Introduction of more observable variables
- Introduction of more actions

Acknowledgement

We thank Prof. Adbul Sattar (Griffith University) and Prof. Ramakoti Sadananda (NICTA) for their support, and the organizers of HAW for the opportunity to present this work.