CAUSAL-DYNAMICS (Improved Understanding of Causal Models in Dynamic Decision Making)

Background

Psychologists have extensively studied dynamic decision-making without formally modelling causality. In contrast, statisticians, computer scientists, and AI researchers have extensively studied causality without considering its central role in human dynamic decision making. This project aligns both fields by building on their respective strengths: providing empirical support for the formal framework where it was previously lacking; and a formal framework for the psychological research.

Our hypothesis is that we can formally model dynamic decision-making from a causal perspective. This enables us to identify both where sub-optimal decisions are made and to identify and recommend the optimal decision. We can usefully test our hypothesis with real world applications involving people making dynamic decisions when interacting with complex systems that are being monitored by devices that reflect the underlying state of the system being controlled or monitored. Such monitored systems include: diabetic-blood sugar monitors, smart meters for tracking home energy use and mobile phone applications for health and life-style tracking.

The consensus from many recent reviews is that the factors influencing the design and use of such devices are poorly understood, resulting in misuse, underuse, and negative consequences. While often overlooked, a key factor shared by all devices, and which plays a major role in the way people make decisions, is that they are *dynamic* in the sense that the user is able to intervene, at any moment, and can change the device and underlying system state that they are monitoring; the information they receive back from the device is also changing, which requires users to have a cognitive causal model of the factors that generate the changes [1]. Users have to intuitively adapt their underlying causal representation of the dynamic elements of the system and change their behaviour in response to its outputs and prompts, such as with blood sugar levels (home diabetic monitoring device) and energy readings (home energy smart meters).

By providing a formal basis for understanding dynamic decision-making the CAUSAL-DYNAMICS project will help us to understand how users make decisions when interacting with devices in order to help them improve the overall outcome. More specifically, optimal change in decision-making behaviour involves trade-offs between needs, motivation and ability, and this is often ignored in the design of devices. Normally, to achieve a goal users require a causal model with the implication that they must learn about the relations between controllable inputs (like consumption of food and drink in the diabetes example, and use of shower, kettle, television in the energy example) and their effects on an outcome (blood sugar level, energy use), while at the same time trading this off against personal needs. Furthermore this challenge is made more complex by the fact that these personal needs will differ depending on whether the goal is ***self-set*** or ***imposed*** [1]. Thus, dynamic decision-making in real world contexts involves trade-offs between motivation/ability to change and optimal change behaviour. CAUSAL-DYNAMICS examines this empirically through a series of cutting-edge research techniques implemented in eight experiments, and incorporated into the Bayesian modelling work adopted in this project.

An emerging theoretical Bayesian framework argues that people use probabilistic causal models in order to reason and make decisions [2, 3, 4, 5, 6]. The principle behind this framework is that our representation of relations in the world, from which we reason and make decisions, assumes that causes change the probabilities of their effects. From this, it is possible to estimate the effects of one's actions on a system, and in turn monitor and control future (or simulate past) actions in the system. Applications of cognitive causal models can be found in many practical domains, such as legal reasoning [7, 8] and systems engineering [9]. We believe that the potential for improved understanding of dynamic decision making can be achieved through the recent transformative work on causal modeling using Bayesian network (BN) approach [2, 10]. Central to that work is the notion of ***idioms*** [9], which are common and repeatable patterns of causal reasoning. However, there is a major research challenge in exploiting this work in the context of improved understanding of self-monitoring devices. While some generic idioms appear to be universally applicable across application domains [2], examples of usable instantiated idioms have so far been restricted to systems engineering and the law; moreover, current BN technology does not support a user-friendly idiom-based approach to modelling. Another research challenge relates to Dynamic Bayesian Networks (DBNs) [11] (which are essentially iterations of a single BN idiom/structure over time). To properly model the dynamic behaviour associated with devices we need to use DBNs, but there are limitations to the current-state-of-the art of DBNs in the sense that they lack a consistent framework for abstraction and efficient implementation.

Thus far, one key limitation in the domain of dynamic control is that there are no formal frameworks from which to establish optimal behaviour. This is particularly important when taking into account what kind of feedback is the most effective for optimal dynamic control in general. Unpredicted or delayed feedback from the outcomes of users' actions with devices leads people to misattribute the outcomes to factors other than their own making, or leads them to ignore the outcome information entirely [12, 13]. However, when people make interventions (by choosing which actions they think will likely generate a desirable outcome – i.e. control as well as monitor) research shows that there is a benefit when it comes to learning about static systems [14, 15, 16, 17]. To date, this has not yet been fully investigated for cases of controlling dynamic systems, despite the ubiquity of dynamic control environments in many areas of practical importance [1, 19].

Two pilot studies [20] conducted in preparation for the CAUSAL-DYNAMICS project provided a roadmap for the research for this proposal. The studies were illuminating because they revealed how varied people's causal models are when making dynamic decisions using a device (home energy smart meters), and what type of measurement tools validly and reliably allow us to understand the causal structures from which real world decisions are built. The proposed research in CAUSAL-DYNAMICS aims to extend the ideas in the pilot studies to improve dynamic decision-making when interacting with any type of device. The data from these studies will be available as part of the BN model validation.

Research hypothesis and objectives

Our main research hypothesis is that when users interact with a device they start with a causal model that is typically different to the optimal causal model. We believe it is possible to address this difference by establishing what users think they are doing, what they are actually doing, and what they should be doing in an optimisation sense. We also believe it is possible to establish whether (and how) people update their causal models in order to understand how the device behaves, and to control both short-term and long-term consequences of their own actions. This behaviour can be assessed against a normative standard provided by dynamic causal models. To address the research hypotheses, the objectives are to:

- Develop a causal (Bayesian) framework (based on a set of causal idioms) that enables researchers to represent: a) how people act; b) how the device works; and c) how people ought to rationally act. This involves extending current approaches to DBNs.
- Conduct a series of simulated experiments in the laboratory to test and improve the causal modelling method and to generate a new dataset that advances our understanding of how users interact with devices.

Meeting the objectives will enable us to:

- Identify ways in which users can make rational decisions in everyday practical situations.
- Combine formal and empirical methods to characterize decision-making processes and the dynamic environment in which they emerge.

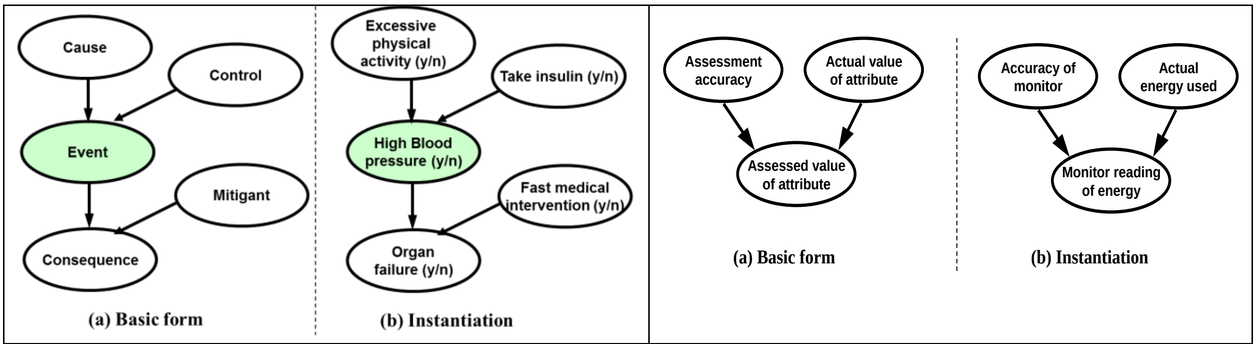
Programme and methodology

The research hypotheses and objectives give rise to the following research questions that drive the programme and methodology (as above we use “device” to refer to any self-monitoring dynamic system).

1. How can we extend the previously proposed causal idioms to adequately capture the core causal reasoning that people use in understanding a device?
2. To what extent do people use causal models in trying to understand devices?
3. To what extent do people re-use common causal idioms to revise their actual causal models?
4. How do people adapt their causal models to cope with a device?
5. How do people cope with trade-off between self-set and imposed goals in their decision making?

There are two main technical work packages to address these questions (WP1: BN modelling/technology; and WP2: Experiments) and a third work package (WP3) dedicated to project management and outreach. WP1 addresses question 1 above and will initially help drive the experiments in WP2 that will address questions 2-5. WP2 will in turn also provide crucial feedback about the specific dynamic idioms required for question 1.

WP1 Bayesian network modelling: This work package will model the actual and ‘ideal’ causal reasoning with BN models. This work will build upon the set of causal idioms identified in [7] and [9], but extend these idioms to deal with dynamically changing environments where agents seek to control the system. Examples of two such idioms are shown in Figure 1.



(a) Joined cause-consequence idiom	(b) Measurement idiom

Figure 1: BN Idioms

One of the benefits of causal representations for learning and inference is that they enable flexible revision in the light of a changing world. This is because generic causal knowledge can be acquired along with the specifics of the system under study, and can be reused/recompiled even when the system undergoes dynamic changes. The modelling expertise provides increasing amounts of ‘formal’ information about the causal models to see the extent to which this knowledge improves the user decision-making/performance. For example, in cases where users initially fail to understand a clear causal relationship, then actually showing them parts of the causal model – maybe even running it for them–should improve their understanding and hence behaviour.

One of the key aims is to develop idioms tailored to the feedback loops characteristic of dynamic systems. To properly model such dynamic behaviour we need to use Dynamic Bayesian Networks (DBNs) [19]. (which are essentially iterations of a single BN idiom/structure over time). As discussed in the background section, limitations to the current-state-of-the art of DBNs means that one of the research challenges we need to overcome in this project is to provide an enhanced notation and associated algorithms to ensure that the full range of causal models identified can be easily specified and executed. Current implementations that support DBNs (for example, in freely available tools such as

AgenaRisk, Genie-Smile and Hugin) are essentially limited to linear groupings of BN components linked by input and output interface nodes; to execute such models the fully expanded underlying models must be used. Since DBNs are essentially iterations of a single BN structure over time, the scale of the model is therefore severely constrained by limitations of existing propagation algorithms.

An objective is therefore to produce extensions to the way DBNs are specified and executed with special focus on the relevant set of idioms. We will develop necessary algorithms and a prototype tool to implement these BN extensions making use of, and extending, existing toolsets and their APIs (Application Programmer Interfaces) where appropriate. These include AgenaRisk, Genie-Smile, Hugin, Matlab, Netica and WinBugs. To support the empirical work, this component of the project will deliver regular prototype tool increments. Using the prototype software, the models will be run using different scenarios as a benchmark of ‘correct’ reasoning and rational decision making, comparing the results from the models with the results from the WP2 experiments. WP1 is divided into five tasks (see Gantt chart for details): 1) review state-of-the-art of BN idioms; 2) DBN idioms specification/notation; 3) DBN algorithms; 4) prototype tool; 5) consolidated idioms based on experimental feedback. There will be a deliverable report for each task (except 4). For task 4 there will be three incremental prototype deliveries. RA1 (under the supervision of Fenton and Neil) will work on these tasks.

WP2 Experiments. The empirical component is driven by eight experiments (labelled E1-E8) that address research questions 2-5. The experiments are laboratory-based decision making tasks. In each case we will base the experiments around two example devices – a smart meter and a medical app. The aim is to use the causal idioms framework to compare optimal behaviour against the kinds of behaviours elicited in the psychological experiments. People will be instructed to learn about and control dynamic systems through the development and use of causal BN models. The data sets generated from this project involve the following: Demographic information, attitudinal data, motivational preferences for devices, choice of interventions, causal diagrams of the system, control-based decision making performance measured against formal model predictions.

Given our review of the literature, we have shown that there are no existing data sets from which to draw on for the purposes of the aims and objectives of this project. Although the empirical work involves combinations of methodologies that are novel, our pilot studies have served three purposes: 1. To determine the feasibility of the empirical design of the experiments, 2. To determine the feasibility of the stimuli and measures incorporated into the experiments, 3. To provide empirical data on a critical first question: The difference between their prior beliefs about their objectives for the device compared to how they actually used it. An illustration of the basic outline of an experiment based on a smart metering device is shown in Figure 2.

Experiments 1-2– Measuring interventional efficiency (addresses Q2-4): E1 will measure planned interventions with frequent feedback, E2 will measure planned interventions following periodic feedback. From this, it is possible to measure the informativeness, or diagnosticity of different intervention choices given the type of feedback they are given.

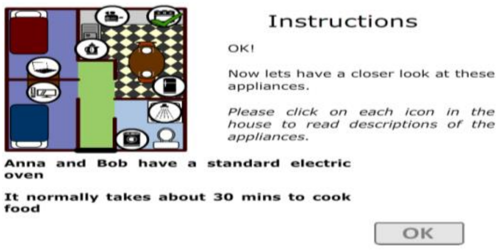
 Screen shot	*Participants observe as an energy monitor reading varies while appliances are switched on and off by the housemates. They are required to intervene on the house by switching appliances on and off, in order to establish the energy consumption of the different appliances. Realistic levels of noise and delays are included in the data stream*.

Figure 2: Basic outline of an experiment

Experiment 3-4– Varying the presentation of information (addresses Q2-4): An important step is to assess the effect of different ways of presenting information on the device and to look at how this impacts on learning to use the device. E3 will involve two different types of outcome feedback times three different Presentation Formats (e.g. different levels of detail and frequency).

Experiment 5-6 – Measuring effective control relative to different goals (addresses Q2-5): We investigate whether users' control of the device is more consistent with self-set goals or imposed goals. In E5 users generate their own self-set goals but will also be presented with imposed goals. There is no explicit instruction about which takes precedent. In E6 we specify which goal must take precedent. Both experiments reveal what preferences they have and what their natural tendency is for controlling the device. This will reveal important information about the potential mismatch between the intended design of the device and how the user benefits, as well as the trade-off decisions made by users.

Experiment 7-8 Sensitivity to changes in the system relative to different goals (addresses Q2-5): An important aspect of dynamic systems is that they have non-stationary properties over time. To examine if users are sensitive to these non-stationary properties we will examine the effectiveness of interventions relative to self-set goals (E7) and imposed goals (E8).

WP2 is divided into 9 tasks (see Gantt chart for full details): one for each experiment (which will each have a deliverable report at the end) and a further task to review the experiments (also resulting in a deliverable report). RA2 will work on these tasks (and will also contribute to WP1, task 5) under the supervision of Osman and Lagnado.

WP3 is divided into two tasks (see Gantt chart [21] for details): project management (which will be undertaken by Fenton) and outreach.

Significance and Dissemination

There will be significant academic impact on both AI research and use, and cognitive science researchers. In AI multiple application domains use BN models for which, in many cases, the construction and implementation of BNs would benefit from the improved modelling support for both BN idioms and DBNs that the project is providing. Current state-of-the-art BN technology has extremely limited or non-existent support for these modelling concepts. The project will deliver open source code so that the academic researchers (as well as users of AI) will be able to use the new modelling and interference techniques in their own work without restriction.

The unique combination of empirical techniques employed will generate original data sets that will advance understanding in the cognitive sciences of how people's causal knowledge informs their decision-making in a dynamic domain. One of the key limitations of research into dynamic decision-making and dynamic causal reasoning is that there are no formal frameworks from which to establish optimal behaviour. Another outstanding contribution that this project will make is through the integration of cutting edge formal modelling techniques that provide precise descriptive and prescriptive representations of the decision-making environment, against which human behaviour can be evaluated. Therefore the programme of combining empirical and formal modelling in this project would substantially advance understanding in the cognitive sciences of dynamic decision-making, providing a causal perspective underpinned by BN models.

The results of both the novel methods and experiments will be published in high-quality refereed journals using the open access route. We will also create a dedicated web page for the project to host all working papers and the prototype software for download.