Project Proposal

“Feature selections by reinforcement learning for bivariate casual inferences of hidden variables”

By

[redacted]

M.Sc. Business Analytics and Intelligence

Master Dissertation Proposal 2019

University of Westminster

Project supervisor

Prof. Thierry Chaussalet

Table of content

Context and Background

Causality

A fully intelligent and autonomous system cannot exclusively rely on preprogramed knowledge. To be able to translate observation into cause-effect relationship will enable human-like knowledge to be computerised on whole new, higher level. The development of casual models has been influent mainly by three branches; computer science, statistics, and econometrics. In the lead of the development within computer science were Juda Pearls. He developed what has become known as the Directed Acrylic Graphs (DAG) approach (Pearl, 2000). From the statistics the potential outcome framework has been developed. It was originally proposed by Jerzy Neyman but have since been further developed (Neyman, 1923). These two models are closely related and have both defined the other as theorems. Econometrics have a few alternatives techniques to look for causality, i.e. controlled regression (including other forms of regressions), instrumental variables, and difference-in-difference (Villa, 2018).

Regardless of approach or applied framework, causality is a difficult and partially philosophical paradox. To determine whether something is the cause of something else could sometimes be common sense but far from always. In fact, it could be the other way around, or it could even be a third, unobserved confounding variable causing both. Due to this, the common belief is that causality cannot be proven unless a controlled experiment is applied. Therefore, obtaining causality from passively observational data is an close to impossible task (Goldthorpe, 1998).

Nonetheless, having knowledge about causality allows for deeper insight into a data set. There are three fundamental analysis which can be explored using causality; probabilistic, interventional, and counterfactual. Probabilistic is an association model determine statistical relationship using i.e. Bayesian Networks and are suitable for decision support (Aliferis, et al., 2003). Inferred causality, also known as hypothetical causality, looks at effects of actions in hypothetical situations (what-if scenarios) by utilising graphic modelling and probabilistic dependency (Pearl, 1998). The last, but arguable the deepest level and most distinct characteristic of causality is counterfactuals. This is based on retro-perspective reasoning and opens for what-if analyses in observational experiments(!). *"Refine the former and interpret the latter"* – Judea Pearl (Pearl, 2000).

Reinforcement Learning

Reinforcement learning methodologies are developed as goal-oriented algorithms which can achieve, through numerous of reinforced steps, complex goals. They are agent based, meaning an agent performs an action in an environment where the state and immediate rewards are being monitored to determine policies and expected long-term values (Sutton & Barto, 2018). It is expected to work well with causality as it is inspired from behaviour psychology. Further, as it is a dynamic programming methodology allows to implement possibilities which cannot yet be defined by a user, making it ideal for causality inferences.

Combination of methodologies

These two methodologies have yet to be fully implemented jointly in the development of a more human-like artificial intelligence. Although this is likely to be achieved in time, the path is not clearly presenting itself and different approaches are being suggested.

Literature review

Reinforcement learning and causality are two important artificial intelligence tools and confines critical educational aspects and knowledge. However, they have until recent years been treated as two separate fields to be studied (Weisswange, et al., 2009). In a paper from 2016, Seng-Beng Ho tries to look at the empirical differences by utilising the two mechanisms for problem solving (Ho, 2016). Ho found that using a casual learning approach allows for faster solutions solving processes, although the optimal solution is less likely to be identified. Reinforced learning on the contrary results in a slower process, but the optimal solution is almost exclusively provided in the approach.

Lake et al. challenged the progress of AI in their 2016 paper and argued for the need of a machine learning which is also able to utilise casual models and have the intelligence of learning to learn (Lake, et al., 2016). Further they provide core components required for such an achievement of human-like reasoning; developmental start-up knowledge, a learning model, understanding of causality structure, and fast inferences interpretation. In addition to this they brought up frequently confronted challenges, such as prior knowledge implementation, biological plausibility in engineering an AI, and lastly a unified casual language. Looking forward they coin the importance of differential programming and 'working memory'.

In chapter 17 of the book Oxford Handbook of Casual Reasoning, M. R. Waldmann takes a psychological approach and looks at the cognitive relationship between reinforced learning and causality and provides evidence of coexistence of the mechanisms in the human brain (Waldmann, 2017). Furthermore, he shows how the brain balance between a model-free and model-based approach to the reinforcement loop. The model-free is less memory expensive but leads to more inaccurate results while the model-based is memory expensive but highly accurate. These are similar characteristics of trade-off which have to be made in the development of a computer algorithm. Towards the end Waldmann rises the issue with latency. The delays in the award and sensory data could throw of the inference of causational states.

In the essay 'From machine learning to machine reasoning' written by Leon Bottou the continuity between algebraic manipulation of knowledge, which could be defined as reasoning, and a large learning system is exploited (Bottou, 2014). Bottou argues for that the need of a complicated 'all-purpose' machine learning system could be circumvented by using logical inferences through manipulations in a learning system.

In a report posted by the Journal of Machine Learning Research Bontempi and Flauder are proposing a new algorithm which makes inferences around causality based on dependency in a multivariate situation. The proposed algorithm, which were given the suitable name D2C, works in a two-step approach, firstly by featurizing relationship through a Bayesian network for then to use a machine learning classifier to learn the occurrence of causal links and directions. Furthermore, they validate they approach through synthetic and published data to determine the methodologies' accuracy. The results are promising and produces significant results in both steps. In addition, Bontempi and Flauder identifies focus areas of future work to be done within feature discovery and alternative assessing of casual links (Bontempi & Flauder, 2015).

The project's aim

Purpose of the project

The intention of this paper is two-folded. The first is to develop a methodology of generating features for the purpose of causal inferences through a reinforced learning environment. The main aim of the paper is focused around the building and requirements of such an environment needed to successfully achieve this. The second is to prove the result of this approach over existent methods with the usage of public available data of the forex market place. Furthermore, by doing that determine causal relationship and links of hidden variables affecting value fluctuations between two currencies.

Justification

The development of the two respective fields of reinforcement leaning and causality, as fields within computer science, have increased rapidly in recent years. As computer power and knowledge are only predicted to improve their paths are inevitably going to cross. This crossing point of artificial intelligence will no doubt lead to more robust and accurate models. It will also concede for a scalability of models surpassing what previously was possible. On the other hand, it would make the availability of computer power more definite and having access to such resources more important. Furthermore, it will endure some weaknesses found in current models of AI such as biases and amount of required data to develop a learning model. Nonetheless, progress is proceeding with constantly new applications and opening of new possibilities for most industries and in most fields of science. However, these could also bring forward some threats, i.e. malicious usage and the fact that it could pilot itself forward without human inputs or understanding, leaving us in fact behind.

The proposed project would lead forward the existing environment, without necessary dealing with the weaknesses addressed in the current situation. Regardless, it is believed that the opportunities are outweighing the threats and a solution of this project would have a positive impact on the status quo. There are few alternative solutions which have been suggested, increasing the bearing on such a model.

The impact of a proposed solution itself would not necessarily be too essential, although as it allows for intervening of two major fields in artificial intelligence most definitely front other prospects and advances of current.

The risk of bringing forward this project is mainly with unfeasibility and unmanageable short-comings in knowledge. Regardless, even this result could prove valuable insight for further progress. To conclude, the validity and justification of the project is defensible and aligned with current trends in the field of artificial intelligence.

The project's objectives

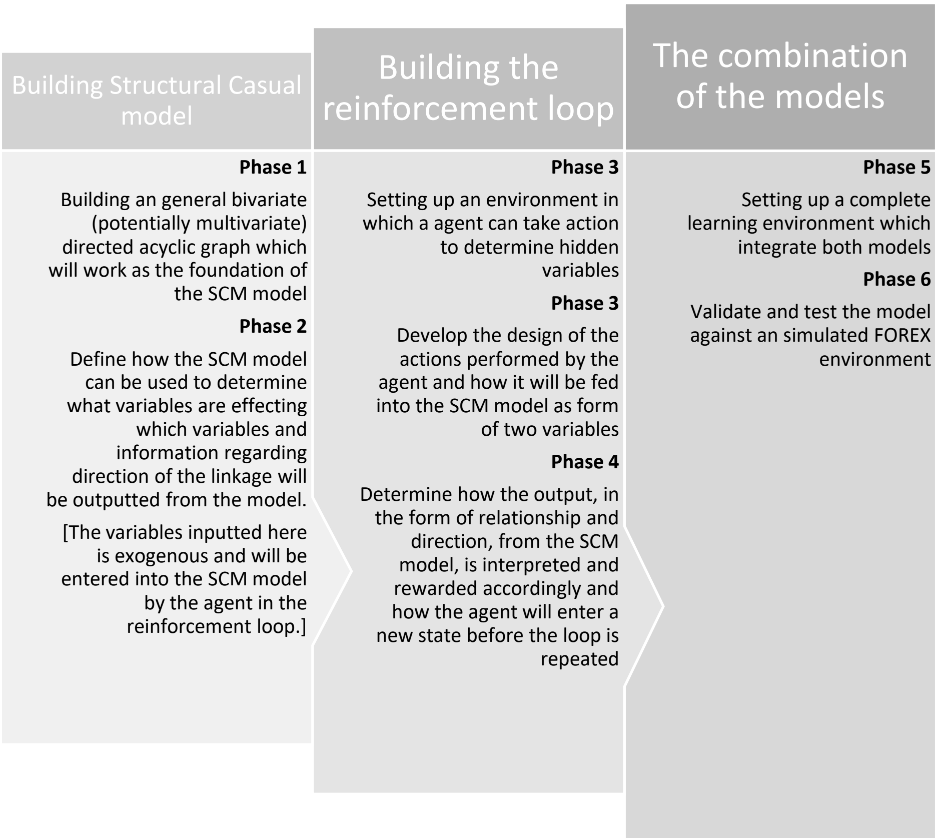
Realisation of the Aim

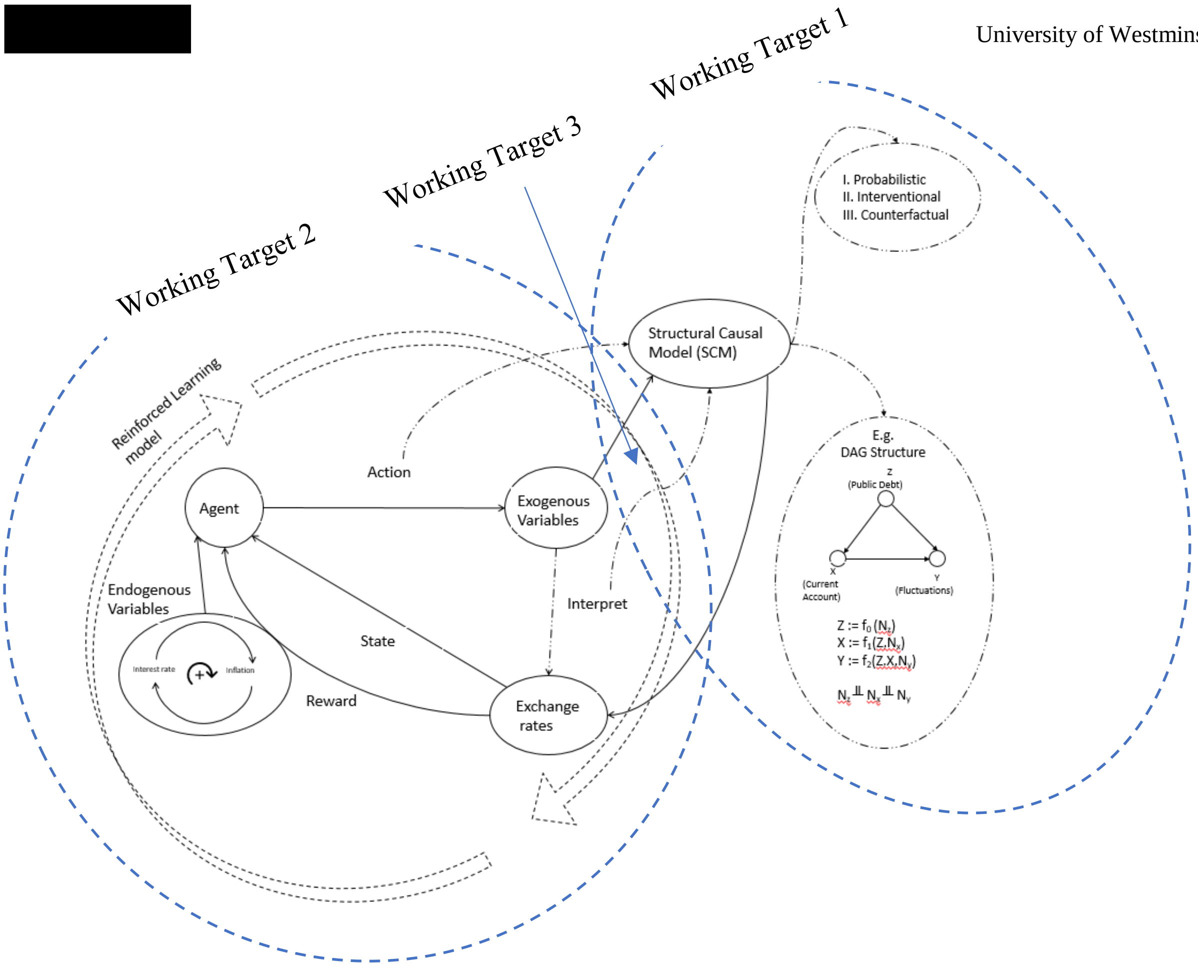
The first objective is to critically asses the current situation of casual model structures and how implementations have been computerised. An alternation of the latest development would then be implemented in the process of the project. The second objective is to develop a combination of reinforcement learning loop and feature creation. The third objective is to combine the two developments into a working model. Here, the causal model developed as the first objective will use the input form the reinforcement learning loop developed as the second objective and provide inputs back into the RL loop, for it to make alternation for the next feature input. Here inputs imply featured causal relationship and direction of linkage, which were previously hidden but 'learned' or discovered by the RL loop. Once these objectives are implementation of data available in the forex market will be used in a simulated environment to verify the model's functionality and accuracy. Any potential output of hidden inferences would then be verified against a real and live environment.

Methodology

Plan

The plan for the project is divided into a six-step proposal, which are covered by three main working targets; designing the structural causal model, building the reinforcement learning loop, and lastly combining the two models in a working environment. All the building and implementation will be done using Python programming language.



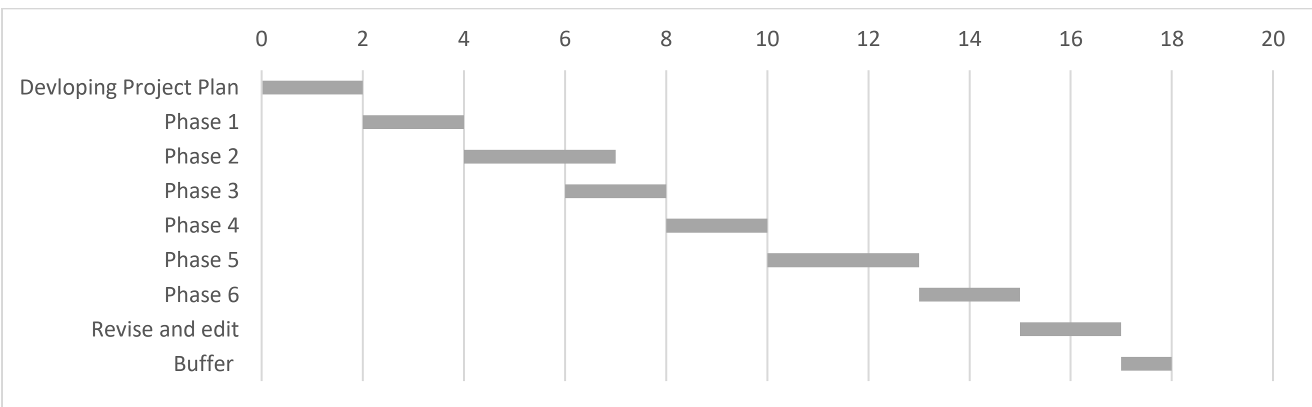


The timeline

The project timeline below, in weeks, starting 1st of May with expected completion 1st of September. Each phase is allocated a working time, with some overlaps as the phases are intervening with each other. Sub-goals have been set to week seven, ten, and fifteen as then two phases creating a working target is completed, and a new part of the project can be initiated.

Each phase will be followed by checking against KPI's. The specific of the KPI's is as following:

- Phase 1: Algebraic formulation of the DAG graph
- Phase 2: Validation of the structural causal model with an accuracy above 90%(?)
- Phase 3: Algebraic formulation and understanding of the reinforcement learning loop
- Phase 4: Validation of the input and output variables against a test set
- Phase 5: Validation of a fully functional, running model
- Phase 6: Validation against a real, live data set



Outcomes

As the project progresses each milestone is expected to produce an outcome. These outcomes are initially required to obtain the same high, standard as is expected in the final product, as these will work as building blocks for the final product. Of this reason, the KPI's previously determined will be monitored closely and are essential as to ensure the expected outcomes are satisfying every requirement. To ensure objective judgment of completed milestone and that they are according to required standard, a small implementation of individual model can be performed and tested out of context.

Criteria's of what is attainable level is based on what standards today methodology achieves. If there is no improvement of the output, in either of the individual model, the alternative model is deemed unsuccessful. The three main outcomes are as mentioned closely linked to the working targets. Hence there are models which they can be compared against separately. The final, real outcome is the setup combining all the three targets. This outcome is the final aim of the project and would of course be more thoroughly tested against multiple of criteria. One of the tests would be the implementation in a simulated forex market. Here, any outputs are yet unknow but it could potentially lead to discoveries of new, hidden variables and relationships.

References

Aliferis, C. F., Tsamardinos, I., Statnikov, A. R. & Brown, L. E., 2003. *Causal Explorer: A Causal Probabilistic Network Learning Toolkit for Biomedical Discovery,* Las Vegas, USA: METMBS.

Bontempi, G. & Flauder, M., 2015. From dependency to causality: a machine learning approach. *Journal of Machine learning Research,* Issue 16, p. 2437.

Bottou, L., 2014. From machine learning to machine reasoning. *Spring Link Machine Learning,* 94(2), p. 133.

Goldthorpe, J. H., 1998. *Causation, Statistics, and Sociology,* Dublin, Irland: The Economic and Social Research Instituit.

Ho, S.-B., 2016. *Causal Learning versus Reinforcement Learning for Knowledge Learning and Problem Solving,* Singapore: Association for the Advancement of Artificial Intelligence.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B. & Gershman, S. J., 2016. *Building Machines That Learn and Think Like People,* NA: Behavioral and Brain Sciences.

Neyman, J., 1923. Neyman–Rubin causal model. *Statistical Science*, p. 463.

Pearl, J., 1998. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* 1st ed. s.l.:Morgan Kaufmann.

Pearl, J., 2000. *Causality.* 2nd ed. New York: Cambridge University Press.

Sutton, R. S. & Barto, A. G., 2018. *Reinforcement Learning: An Introduction.* 2nd ed. Cambridge: MIT Press.

Villa, A. R. d., 2018. *Towards Data Science.* [Online]
Available at: https://towardsdatascience.com/why-do-we-need-causality-in-data-science-aec710da021e
[Accessed 09 04 2019].

Waldmann, M. R., 2017. *Oxford Handbook of Causal Reasoning.* 1st ed. Göttingen, Germany: Oxford University Press.

Weisswange, T. H., Rothkopf, C. A., Rodemann, T. & Triesch, J., 2009. *Can reinforcement learning explain the development of causal inference in multisensory integration?,* Shanghai, China: IEEE 8th International Conferance.