Causal Effects for Prediction and Deliberative Decision Making of Embodied Systems

Nihat Ay
Keyan Zahedi

SFI WORKING PAPER: 2011-11-055

Causal Effects for Prediction and Deliberative Decision Making of Embodied Systems

Nihat Ay[1,2] and Keyan Zahedi[1]
[1]MPI for Mathematics in the Sciences, Inselstrasse 22, 04103 Leipzig Germany
[2]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA
e-mail: {nay,zahedi}@mis.mpg.de

Summary

This article deals with the causal structure of an agent's sensori-motor loop and its relation to deliberative decision making. Of particular interest are causal effects that can be identified from an agent-centric perspective based on in situ observations. Within this identification, an optimal world model of the agent plays a central role. Its optimality is characterized in terms of prediction quality.

1 Introduction

Evaluating different possibilities and deliberatively choosing among them is an important ability in humans and animals. In order to be intentional, such a choice has to be based on knowledge about causal consequences of individual brain dynamics. Within dynamical systems theory, a plausible model to describe switching between different dynamics is based on chaotic attractors [2]. However, in this framework the study of causality remains a challenge. In this paper, we address causal effects in the sensori-motor loop (SML) within a coarse-grained level of description where transitions are modelled in terms of stochastic maps.

We use the formalism of Bayesian networks to study the causal relations in the SML (see previous work [3] and [1] in this context). A Bayesian network consists of two components, a directed acyclic graph Γ and a set of stochastic maps describing the individual mechanisms of the nodes in the graph. More precisely, Γ is assumed to have no directed cycles (see Fig. 1 as an example). Given a node Y with state set $\mathscr{Y}$, we write $X := pa(Y)$ for the set of nodes X_i with an edge from X_i to Y. The mechanism of Y is formalized in terms of a stochastic map $\kappa : \mathscr{X} \times \mathscr{Y} \to [0,1]$, $(x,y) \mapsto \kappa(x;y)$, where $\mathscr{X}$ denotes the state set of X. The stochasticity of κ refers to $\sum_y \kappa(x;y) = 1$ for all x.

The Fig. 1 shows the general causal diagram for the SML, where W_t, S_t, C_t, A_t denote the world, sensor, controller (memory), and action at some time t. We denote their state sets by $\mathscr{W}, \mathscr{S}, \mathscr{C}, \mathscr{A}$, respectively. The stochastic maps α, β, φ, and π describe the mechanisms that are involved in the sensori-motor dynamics. Here, φ and π are intrinsic to the agent. They are assumed to be modifiable in terms of a learning process. The mechanisms α and β are extrinsic and encode the agent's embodiment which sets constraints for the agent's learning (for details, see [5, 3]).

Pearl [4] proposes the concept of intervention to capture causal relationships between random variables in a given Bayesian network. We will show that the formalization of the SML allows to determine causal relations solely observational without any experimental intervention, although its derivation is based on the concept of intervention (see Sec. 2). In this identification of causal effects, the optimal world model plays a central role. It is given as the conditional probability $p(s|c,a)$ of observing the next sensor state s as a result of the current controller state c and the current action a of the agent.

2 Causal effects in the Sensori-Motor Loop

Fig. 1 illustrates the causal structure of the SML. This representation has been used in [3, 1].

Pearl's formalism [4] allows to define and study causal effects in the SML, for instance the effect of actions on sensor inputs. Here, a fundamental understanding is that in order to reveal causal effects one has to test the system in experimental situations (ex situ). In this context, intervention is an operation that serves as an important build-

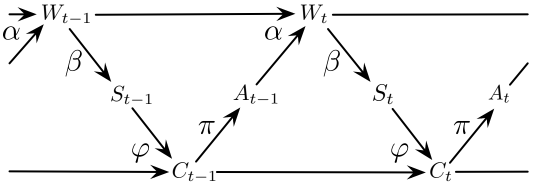

Figure 1: Causal diagram of the sensori-motor loop.

ing block in corresponding experiments. However, it is not always possible for an agent to perform an intervention. Therefore, it is important to know whether a particular causal effect can be identified purely based on in situ observations of the agent. In the proposition below, we list three causal effects in the SML that are identifiable by the agent without actual intervention. In order to be more precise, we have a closer look at the causal diagram of the transition from time $t-1$ to t.

Here, as shown in Fig. 2, we consider the future sensor value of only one time step and summarize the past process by a varible H_{t-1}. We focus on the resulting causal diagram of Fig. 3 (left-hand side). The joint distribution in the reduced diagram is given as

$$\begin{aligned} & p(h, c, a, w, s) \\ &= p(h)\varphi(h; c)\pi(c; a)\alpha(h, a; w)\beta(w; s). \quad (1) \end{aligned}$$

Given such a factorization of the joint distribution, one can define the intervention in a subset, which is referred to as *do*-operation. It is simply given by the corresponding truncation of the product, which formalizes the idea that the mechanisms of the intervened variables are changed from outside. As an illustration, we consider the product (1) and set the value A to a, that is we do a. The result of

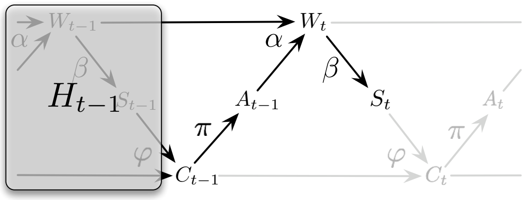

Figure 2: Reduction procedure of the causal diagram.

this conditioning is given as

$$p(h, c, w, s|do(a)) = p(h)\varphi(h; c)\alpha(h, a; w)\beta(w; s).$$

Summing over the variables h, c, w, for example, gives us the probability of observing s after having set a. The corresponding stochastic map is referred to as the *causal effect* of A on S:

$$p(s \,|\, do(a)) \;=\; \sum_{h,c,w} p(h, c, w, s \,|\, do(a)).$$

Note that, in general, we do *not* have $p(s \,|\, do(a)) = p(s \,|\, a)$, which is an important property of causal effects. Applying the described procedure, one can compute various other causal effects. The following question plays a central role in Pearl's causality theory: Is it possible for an observer, such as an autonomous agent considered in this paper, to reveal a causal effect based on observations only? At first sight, this so-called identifiability problem appears meaningless, because causal effects are based on the concept of intervention. However, having some structural information sometimes allows to identify causal effects from observational data.
The following causal effects can be identified by the agent without any actual intervention.

Proposition 1. *Let the joint distribution (1) be strictly positive. Then the following equalities hold:*

(**a**) $p(s \,|\, do(a), c) := \dfrac{p(s, c \,|\, do(a))}{p(c \,|\, do(a))} = p(s \,|\, c, a)$

(**b**) $p(s \,|\, do(a)) \;=\; \sum_{c} p(s \,|\, c, a)\, p(c)$

(**c**) $p(s \,|\, do(c)) \;=\; \sum_{a} p(a \,|\, c) \sum_{c'} p(s \,|\, c', a)\, p(c')$.

The proof of Proposition 1 is given in the appendix. In all three causal effects of this proposition, the conditional distribution $p(s \,|\, c, a)$ turns

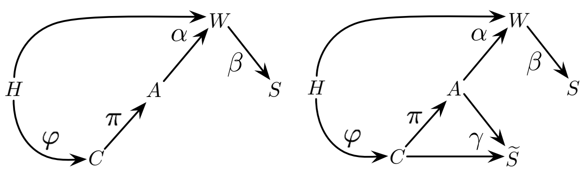

Figure 3: Left: Reduced causal diagram for one time step. Right: Causal diagram with world model γ.

out to be essential as building block for the identification of the causal effects. Note that in the strictly positive case, according to Proposition 1 (a), it is not dependent on the agent's policy. In the next section, this distribution will be studied in more detail.

3 World model & prediction

The causal effects of Proposition 1 involve the conditional distribution $p(s \mid c, a)$. In this section we derive an interpretation of this conditional distribution as optimal world model that allows for the best possible prediction. In order to do so, we extend the causal diagram of Fig. 3 by a world model γ which assigns a probability of observing s as a result of the action a in the context of the internal state c, formally $\gamma : \big(\mathscr{C} \times \mathscr{A}\big) \times \mathscr{S} \to [0,1]$. The world model is a model of the agent's expectation, which can be used for a prediction $\widetilde{S}$ of the next sensor input S. We obtain the diagram of Fig. 3 (right-hand side).

The distribution of $\widetilde{S}$ given C is derived as

$$\begin{aligned} \widetilde{p}(s|c) &:= \mathsf{Prob}\{\widetilde{S} = s \mid C = c\} \\ &= \sum_{a'} \pi(c; a')\, \gamma(c, a'; s). \end{aligned}$$

(Here, Prob stands for probability.) In order to measure the quality of the world model γ, we use the entropic distance, also known as KL divergence, between $\widetilde{p}(s|c)$ and $\beta(w; s)$:

$$D(\beta \,\|\, \widetilde{p}) \;:=\; \sum_{c,w} p(c,w) \sum_{s} \beta(w;s) \ln \frac{\beta(w;s)}{\widetilde{p}(s|c)}.$$

The following proposition identifies the conditional probability $p(s|c,a)$ as best world model in terms of this deviation measure.

Proposition 2. *If a world model $\widehat{\gamma}$ satisfies $\widehat{\gamma}(c,a;s) = p(s|c,a)$ whenever $p(c,a) > 0$ then it minimizes the distance $D(\beta \,\|\, \widetilde{p})$:*

$$\begin{aligned} &\inf_{\gamma} \sum_{c,w} p(c,w) \sum_{s} \beta(w;s) \ln \frac{\beta(w;s)}{\sum_{a'} \pi(c;a')\, \gamma(c,a';s)} \\ &= \sum_{c,w} p(c,w) \sum_{s} \beta(w;s) \ln \frac{\beta(w;s)}{\sum_{a'} \pi(c;a')\, p(s|c,a')}. \end{aligned}$$

This implies that the minimal distance coincides with the conditional mutual information $I(W; S \mid C)$.

The proof of Proposition 2 is a straightforward application of the Lagrange multiplier method.

4 Deliberative actions

In the previous sections, we referred to a reactive interpretation of actions (see [4], page 108). The corresponding reactive policy π assigns an action a to the agent's state c, formalized in terms of a stochastic map $\pi : \mathscr{C} \times \mathscr{A} \to [0,1]$, $(c,a) \mapsto \pi(c;a)$. In order to have a deliberative interpretation of a given policy, we decompose the internal state c into a memory state m and a goal state g, that is $c = (m,g)$ and $\mathscr{C} = \mathscr{M} \times \mathscr{G}$. We assume that the goal variable G only depends on the memory variable M, which leads to the causal diagram of Fig. 4.

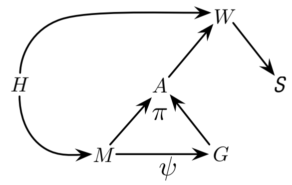

Figure 4: Causal diagram with goal map ψ and deliberative policy π.

The map ψ models the process of choosing a goal. Studying intentionality, we hypothesize that the causal effect $p(s \mid do(g))$ of the goal g on the sensor input s plays an important role. A simple calculation, similar to the one in the proof of Proposition 1 (b), yields

$$\begin{aligned} &p(s \mid do(g)) \\ &\quad= \sum_{m} p(s \mid m, g)\, p(m) \\ &\quad= \sum_{m} \sum_{a} p(s \mid m, a)\, \pi(m, g; a)\, p(m). \end{aligned}$$

In particular, the causal effect of the goal variable G on the sensor variable S is identifiable. This implies that intentionality is identifiable through M, G, A, S, a conceptually interesting observation. Note, however, that this is only valid in the case of strict positivity of the underlying distribution.

Acknowledgements

Both authors thank Ralf Der, Daniel Polani, and Bastian Steudel for valuable discussions on causal effects in the sensori-motor loop. This work has been supported by the Santa Fe Institute.

References

[1] N. Ay and D. Polani. Information flows in causal networks. *Advances in Complex Systems*, 1(11):17–41, 2008.

[2] K. Kaneko and I. Tsuda. *Complex Systems: Chaos and Beyond*. Springer, 2001.

[3] A. S. Klyubin, D. Polani, and C. L. Nehaniv. Tracking information flow through the environment: Simple cases of stigmergy. *Proceedings of Artificial Life IX*, 2004.

[4] J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.

[5] K. Zahedi, N. Ay, and R. Der. Higher coordination with less control – a result of information maximization in the sensori-motor loop. *Adaptive Behavior*, 18(3–4):338–355, 2010.

Appendix

Proof of Proposition 1:

(a) $$p(h,c,w,s\,|\,do(a)) = p(h)\,\varphi(h;c)\,\alpha(h,a;w)\,\beta(w;s).$$

This implies

$$p(s,c\,|\,do(a)) = \sum_{h,w} p(h)\,\varphi(h;c)\,\alpha(h,a;w)\,\beta(w;s)$$

$$p(c\,|\,do(a)) = \sum_{s}\sum_{h,w} p(h)\,\varphi(h;c)\,\alpha(h,a;w)\,\beta(w;s) = p(c)$$

$$\begin{aligned}
p(s\,|\,do(a),c) &= \frac{p(s,c\,|\,do(a))}{p(c\,|\,do(a))} \\
&= \sum_{h,w} \tfrac{p(h)}{p(c)}\,\varphi(h;c)\,\alpha(h,a;w)\,\beta(w;s) \\
&= \sum_{h,w} p(h\,|\,c)\,p(w\,|\,h,a)\,p(s\,|\,w) \\
&= \sum_{h,w} p(h\,|\,c,a)\,p(w\,|\,h,a,c)\,p(s\,|\,w,h,a,c) \\
&\quad \text{(conditional independence, see diagram in Figure 3)} \\
&= p(s\,|\,a,c).
\end{aligned}$$

The second and third equations of the proposition follow from the general theory (see [4], Theorem 3.2.2 (Adjustment for Direct Causes, and Theorem 3.3.4 (Front-Door Adjustment)). For completeness, we prove them directly.

(b) $$\begin{aligned}
p(s\,|\,do(a)) &= \sum_{h,c,w} p(h,c,w,s\,|\,do(a)) \\
&= \sum_{h,c,w} p(h)\,\varphi(h;c)\,\frac{\pi(c;a)}{p(a|c)}\,\alpha(h,a;w)\,\beta(w;s) \\
&= \sum_{h,c,w} \frac{p(h,c,a,w,s)}{p(c,a)}\,p(c) \\
&= \sum_{c} p(s|c,a)\,p(c).
\end{aligned}$$

(c) $$\begin{aligned}
p(s\,|\,do(c)) &= \sum_{h,a,w} p(h,a,w,s\,|\,do(c)) \\
&= \sum_{a} \pi(c;a) \sum_{h,w} p(h)\,\alpha(h,a;w)\,\beta(w;s) \\
&= \sum_{a} p(a|c) \sum_{h,w} \left(\sum_{c'} p(c')\,p(h|c')\right) \times p(w|h,a)\,p(s|w) \\
&= \sum_{a} p(a|c) \sum_{c'} p(c') \times \sum_{h,w} p(h|c')\,p(w|h,a)\,p(s|w) \\
&= \sum_{a} p(a|c) \sum_{c'} p(c') \times \sum_{h,w} p(h|c',a)\,p(w|h,a,c')\,p(s|w) \\
&= \sum_{a} p(a|c) \sum_{c'} p(c')\,p(s|c',a).
\end{aligned}$$

□