Identification in Chain Multi-Agent Causal Models

Sam Maes and **Stijn Meganck** and **Bernard Manderick**
Vrije Universiteit Brussel
Computational Modeling Lab (CoMo)
Departement Informatica
Pleinlaan 2, 1050 Brussels
Belgium

Introduction

In this paper we introduce chain multi-agent causal models which are an extension of causal Bayesian networks to a multi-agent setting. Instead of 1 single agent modeling the entire domain, there are several agents organised in a chain, each modeling non-disjoint subsets of the domain. Every agent has a causal model over the variables in his domain, determined by an acyclic causal diagram and a joint probability distribution over its observed variables. See Figure 1 for an example.

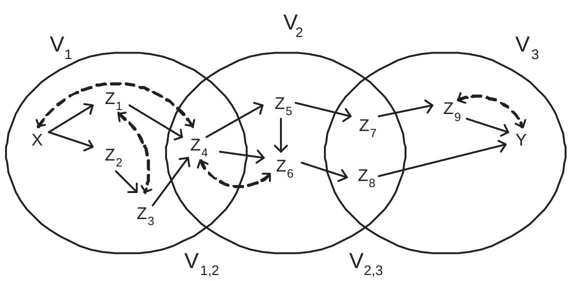

Figure 1: A 3-agent chain model.

We study the identification of causal effects, which is the calculation of the effect $P_x(s)$ of manipulating a variable X on other variables S from observational data and an acyclic causal diagram, containing both arrows and bi-directed arcs. An arrow indicates a direct causal relationship between the corresponding variables from cause to effect, meaning that in the underlying domain there is a stochastic process *P(effect|cause)* specifying how the effect is determined by its cause. Furthermore this stochastic process must be autonomous, i.e., changes or interventions in *P(effect|cause)* may not directly influence the assignment of other stochastic processes in the domain. A bi-directed arc represents spurious dependencies due to unmeasured confounders, this is an unobserved common cause of the corresponding variables.

Deciding if a causal effect is identifiable amounts to assessing whether the assumptions of a diagram are sufficient to calculate the effect of the desired intervention from observational data. When all variables of a domain can be observed, all causal effects are identifiable. In the presence of unmeasured confounders, identifiability becomes an issue. See (Pearl 2000) and (Tian & Pearl 2002) for treatment of this problem in single agent causal models.

In this work we extend an existing single agent identification algorithm to chain multi-agent causal models. Given some assumptions, we provide a technique to calculate the effect of manipulating a variable in agent *A* on some variables in another agent *B*, while only communicating information concerning variables that are shared between neighboring agents on the chain between *A* and *B* and variables that are being studied in that specific query.

The main advantages of the multi-agent solution is that the identification of causal effects can be assessed without disclosing sensitive information of a local model to other agents. It allows to perform causal inference in situations where parts of the model are kept confidential by their distributors.

Imagine for example a chain bi-agent causal model, where 2 credit card companies want to assess the effect of changing the debet limit on the amount of fraudulous transactions being performed. One company could have a model over credit card users which relates the debet limit of a user to its general credit card usage, such as: average amount of monthly transactions, average amount of each transaction, etc. The other company could have a model over the credit card terminal owners (shops), relating the type, geographical location, time of year, etc. of the shop to the amount of fraudulous transactions performed in each shop.

With the techniques introduced in this paper the 2 companies could use the information stored in both their models to calculate the wanted effect, while only communicating over shared variables.

Chain Multi-Agent Causal Models

A chain multi-agent causal model (CMACM) consists of n agents, each of which is represented by $M_i = \langle V_i, G_i, P(V_i), (K_i^-, K_i^+) \rangle$.

- V_i is the subset of variables agent-i can access.
- G_i is the causal diagram over variables V_i.
- $P(V_i)$ is the joint probability distribution over V_i.

- (K_i^-, K_i^+) stores the intersections with neighboring agents on the chain $V_{i-1,i} = \{V_i \cap V_{i-1}\}$ and $V_{i,i+1} = \{V_i \cap V_{i+1}\}$ respectively. We assume that the agents agree on the structure and the distribution of their intersections.

  In the CMACM of Figure 1:

$$\begin{aligned} V_1 = \{X, Z_1, Z_2, Z_3, Z_4\},\ \ V_{1,2} = \{Z_4\} \\ V_2 = \{Z_4, Z_5, Z_6, Z_7, Z_8\},\ \ V_{2,3} = \{Z_7, Z_8\} \\ V_3 = \{Z_7, Z_8, Z_9, Y\} \end{aligned}$$

In (Maes, Meganck, & Manderick 2004; 2005) we extended a single agent algorithm to the bi-agent case, i.e., a setting where there are 2 agents each modeling a part of the domain. There is 1 agent with the model that contains the manipulated variable X and the other with the model containing the variable(s) S to be studied. Some of the assumptions for the algorithm to calculate $P_x(s)$ are:

1. there is no bi-directed edge connected to any child of X.
2. $V_2 \perp\!\!\!\perp V_1 | V_{1,2}$, i.e., V_2 is conditionally independent of V_1 given the intersection.
3. $Pa(Ch(X)) \subset (V_1 \cup V_{1,2})$.

Chain Multi-agent Identification

First we introduce the notation $P_x^V(d)$, meaning that $P_x(d)$ is calculated in the model with variables V. When the superscript is dropped, this implies that the entire set of variables is being used. We start with the following lemma:

Lemma 0.1 (Domain Reduction)
Consider a causal model with variables V, consisting of disjoint subsets A, B, and D, and $V = (A \cup B \cup D)$. Let $X \in A$ and $P_x(d)$ be identifiable using a single agent identification algorithm, with $Ch(X) \cup Pa(Ch(X)) \subset (A \cup D)$. Then, the following holds:

$$P_x^V(d) = P_x^{V \backslash B}(d) = P_x^{A,D}(d) \tag{1}$$

This lemma implies that if no bi-directed edge is connected to any child of X, $P_x(d)$ can be calculated equivalently in the entire domain V and in $V \backslash B$, where B is a set of variables which contains no children of X and parents of children of X.

Next, we introduce a lemma that combines Lemma 0.1 and bi-agent identification from (Maes, Meganck, & Manderick 2004).

Lemma 0.2 (Recursive Chain Identification)
Consider a setting where D is a D-separation set between X, X-rest and Y, Y-rest. If the assumptions of bi-agent identification also hold with $V_1 = (X, X\text{-rest})$, $V_{1,2} = D$ and $V_2 = (Y, Y\text{-rest})$, then

$$P_x(y) = \sum_D \left(P_x^{X, X\text{-rest}, D}(d) . \sum_{Y\text{-rest}} P(y, y\text{-rest}|d) \right) \tag{2}$$

This lemma implies that if D is a separation set, the calculation of $P_x(y)$ can be separated in 2 parts: first calculating the effect of X on D in the model $(X, X\text{-rest}, D)$, then multiplying this with the sum over Y-*rest* of $P(y, y\text{-rest}|d)$.

Theorem 0.3 (Multi-Agent Chain Identification)
Consider a setting as in Figure 1, where $V_{2,3}$ is a D-separation set between V_3 and the remaining variables. Then:

$$P_x(s) = \sum_{V_{2,3}} \left(P_x^{V_1, V_{1,2}, V_2, V_{2,3}}(v_{2,3}) . \sum_{V_3 \backslash S} P(v_3 | v_{2,3}) \right) \tag{3}$$

$P_x^{V_1, V_{1,2}, V_2, V_{2,3}}(v_{2,3})$ will be calculated by applying bi-agent identification, where the calculations over $(V_1, V_{1,2})$ and $(V_{1,2}, V_2, V_{2,3})$ are performed separately. After the part over $(V_1, V_{1,2})$ is marginalized to a distribution over $(X, V_{1,2})$, it is communicated to agent2, where the rest of the calculation is performed.

This theorem implies that multi-agent identification can be performed over a chain of 3 agents, starting with a calculation between the first two agents, resulting in a distribution over X, the variable that is being intervened on, and $V_{2,3}$, the intersection between agent2 and agent3. Separately, in agent3 a local calculation is performed which results in a distribution over $V_{2,3}$, i.e. the intersection with agent2, and S, the set of variables that we are studying. This result can easily be extended to chains with > 3 agents, by iteratively applying the theorem. The proofs of these lemmas and theorem are out of scope for this paper.

Conclusion

In this paper we introduced the paradigm of chain multi-agent causal models (CMACMs), which are an extension of causal Bayesian networks to a multi-agent setting, where the agents are organised in a chain, so that an agent has a maximum of 2 neighbors. x Given some assumptions, we provided an exact technique to calculate the effect of manipulating a variable in agent *A* on some variables in another agent *B*, while only communicating information concerning variables that are being shared between neighboring agents on the path between agents *A* and *B* and variables that are being studied in that specific query (i.e. X and Y in $P_x(y)$).

References

Maes, S.; Meganck, S.; and Manderick, B. 2004. Multi-agent identification of causal effects. In Ghidini, C.; Giorgini, P.; and van der Hoek, W., eds., *Proceedings of the 2004 European Workshop on Multi-Agent Systems (EUMAS)*, 409–417.

Maes, S.; Meganck, S.; and Manderick, B. 2005. Causal inference in multi-agent causal models. In *Proceedings of Modèles Graphiques Probabilistes pour la Modélisation des Connaissances, Atelier of EGC 05*, 53–62.

Pearl, J. 2000. *Causality: Models, Reasoning and Inference*. MIT Press.

Tian, J., and Pearl, J. 2002. On the identification of causal effects. Technical Report (R-290-L), UCLA C.S. Lab.