# A Proposal on Discovering Causal Structures in Technical Systems by Means of Interventions

*Josephine Rehak*

Vision and Fusion Laboratory
Institute for Anthropomatics
Karlsruhe Institute of Technology (KIT), Germany
Josephine.Rehak@kit.edu

## Abstract

Causal Discovery has become an area of high interest for researchers. It has lead to great advances in medicine, in the social sciences and in genetics. But up til now, it is hardly used to identify causal relations in technical systems. This paper presents the basic building blocks for in-depth research. This paper reviews established causal discovery methods and causal models. In contrast to existing surveys of this domain, we focus on the causal discovery methods using interventions. Based thereon, we propose the idea of a promising interventional discovery approach for technical systems. It takes advantage of not only direct, but also indirect causal relationships, which might improve the learning process of causal structures.

## 1 Introduction

In 360 B.C., Plato already knew that everything in the universe had a cause and is thus an effect of that cause. Nowadays, causality has emerged as the most fundamental common theme in the sciences. Scientists from domains as oceanography, genetics, philosophy or psychology research on discovering causal relationships. Researching causality has strong prospects for the future, because it can improve human-level artificial intelligence, medicine, precision agriculture or production-optimized factories. In general, learning about causal relations helps to understand technical as well as social systems. By using it, one can better predict possible behavior, alternative outcomes and create policies. Naturally, each domain has its own characteristics. In the social sciences, experiments are rather expensive and are limited by social ethics. Fortunately, we have broader opportunities in the domain of technical systems. They offer the advantage of systematic investigations, while no human lives are at risk. But they also come with new hurdles: Some interventions may not be possible, because a change at the place of investigation is not feasible or could damage the system. We need an approach, which respects these boundaries. The method should efficiently identify causal relationships using a minimal amount of data. In the best case, we can make use of existing algorithms to achieve high knowledge gain with minimal costs. To identify such approaches, we shed light on established methods and causal models. By presenting the current state of research, we join a series of publications such as [7, 8, 13].

## 2 The Nature of Causal Relations

Causal relations tend to be complex. The most common definition of causality is, that given two events $A$ and $B$, $A$ causes $B$, when $B$ relies on $A$ for its value. This causal relation of $A$ and $B$ implies several properties. One is the time-dependency: the event $B$ must not happen before $A$. $B$ might occur after event $A$, but the delay time, also called dead time, between $A$ and $B$ must be adhered to. Further on, a causal relation always implies a correlation, but a correlation does not imply a causation. In contrast to correlation, the causal relation guarantees a joint appearance of events $A$ and $B$. Also the deletion of the cause $A$, always implies the deletion of $B$, assumed no other event causes $B$. Besides the discussed relation $A \rightarrow B$, in the deterministic world there exist four other possible causal relations between the two variables $A$ and $B$. The easiest one is $A \leftarrow B$, where $A$ depends on $B$ for its value. It is as $A \rightarrow B$ a simple directed relation and we assume them to be the most common relations in nature.

Another possible relation is causal independence denoted as $A \perp B$. In this case, does neither A depend on B, nor B depend on A. It might occur, that the two variables correlate, what is then called a spurious correlation. The trickiest causal relations between two variables are bi-directed causal relations and confounding relations. In a bi-directed relation $A$ and $B$ causally depend on each other. In a confounding relation, $A$ and $B$ are caused by an unmonitored third variable and are indirectly causally related. The whole domain of confounder analysis is solely devoted to finding such relations. Both cases are difficult to detect as they are hard to keep apart from each other and can easily misidentifies as $A \rightarrow B$ or $A \leftarrow B$. Confounding and bi-directed relations have been be denoted as $A \leftrightarrow B$ in the literature. Next to this, there also exists the notation of $A - B$ for bi-directed relations. We will stick to the later in this report.
These were all the possible relations between two variables, but a causal graph with more variables can be much more complex. A variable might actually depend on multiple variables for its value. Because of this, the number of possible models grows exponentially with the number of variables. To reduce this amount of possible models is the main goal of causal discovery methods. In the best case, the true underlying causal model can be uncovered.

# 3 An Overview of Causal Models

A causal model describes the causal dependencies in a system or population. Causal models can represent only a small part of reality. As every cause has its own cause, the causal network of reality is too gigantic to be fully captured. Hence, a causal model always implies a trade off between complexity and completeness. In literature, the events on the boundary of a model are called exogenous variables. Their value stems from outside of the system. The cause of an event inside the model is called an endogenous event. Since causal models replicate reality, in reverse reality can falsify the models. Different types of causal models have been developed and their number is rising. For these reasons, we limit ourselves to the most popular causal models.

## 3.1 Structural Causal Models

Structural Causal models (SCMs) were introduced by [24]. They consist of a set of exogenous variables $V$ an a set of endogenous variables $W$ and a set of structural equations $F$. SCMs belong to the family of Structural Equation Models (SEMs), also called Functional Causal Models (FCM). These kind of models use equations to represent the graph structure. The definition of SEM and SCM is ambiguous. Sometimes SCMs are referred as SEMs. By our definition, SEMs use linear equations, while SCMs use functions. Consequently, SCMs are more powerful than SEMs. In both models, each equation contains an independent noise variable $U$, which contain the patterns that cannot be causally explained. The probabilistic extension of SCMs are Bayesian Networks.

## 3.2 Bayesian Networks

Bayesian Networks (BNs), or Belief Networks, belong to the family of Probabilistic Graphical Models (PGMs). They combine graph structures with (conditional) probability distributions. The graph $G$ of a BN consists of a set of random variables $X$ and a set of directed edges $E$ connecting the variables. Per definition, the graph is a directed acyclic graph (DAG), as the edges must not form cycles. Each variable has a probability distribution assigned. According to the Law of Total Probability, all probabilities in a probability distribution sum to one [26]. Popular enhancements of BNs are Dynamic Bayesian Networks (DBNs) [5], Object-Oriented Bayesian Networks (OOBNs) [15].

## 3.3 Markov Random Fields

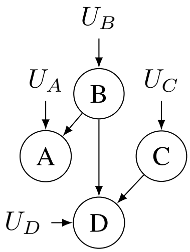
Another way or presenting causal relations are Markov Random Fields [14]. As a PGM, MRFs consist of a set of variables $V$ and a set of probability distributions $P$. Here, the edge weights do not sum to one. To receive normalized probabilities, we have use the normalization constant $Z_\phi$. In opposite to BNs, MRFs contain bidirected edges, which may form cycles. But unidirected relations, as in BNs, are no allowed. Further developments of MRFs are Gaussian Markov Random Fields [29] or Hidden Markov Random Fields [18].

### 3.4 Acyclic Directed Mixed Graphs

Acyclic Directed Mixed Graphs (ADMGs) were introduced by Judea Pearl [12]. In their essence, ADMGs consist of multiple Bayesian Networks and thus allow bidirected edges. Their only constraint is that unidirected edges are not allowed to form cycles. The bidirected edges stand for a latent common cause, which is not included in the variable set. Successors of the ADMGs are alternative Acyclic Directed Mixed Graphs (aADMGs) [27].

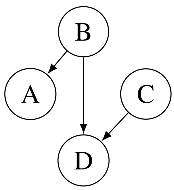
## 4 Methods for Causal Discovery

The purpose of causal discovery is to recreate the underlying true causal model $G*$ from the set of possible models. This can be done by an algorithm or user posing queries to the data. In observational causal discovery, the query concerns the causal relationship between two variables. A causal discovery method then tries to answer the query using prerecorded data. In interventional causal discovery, the query often concerns the question how the relation between two variables changes, when we intervene on one of them. In this case, the discovery method needs access to a live system which reacts to its interventions. To perform causal discovery, one expects that the Causal Markov assumption holds. This asserts, that the data is generated by an underlying model $G*$ and not by chance. The use of experiments with interventions is the oldest causal discovery approach. Back in 1982 Paul Holland stated: "No Causation without manipulation". He identified interventions as the only method to discover causal relations [10]. Contradicting to this, one could observe a trend to observational causal discovery in the nineties. This trend was crucial for domains where experiments come at high cost and effort. In recent years, there has been a trend towards methods handling a mixture of observational data and data from interventions [22].



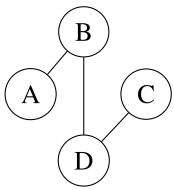

$$
\begin{aligned}
V &= \{A, D\} \quad \textit{endogenous Variables} \\
W &= \{B, C\} \quad \textit{exogenous Variables} \\
F &= \{A = f_A(B, U_A), \quad \textit{Structure Equations} \\
&\quad B = f_B(U_B), \\
&\quad C = f_C(U_c), \\
&\quad D = f_D(C, B, U_D)\}
\end{aligned}
$$

(a) Structural Causal Model



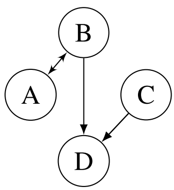

$$V = \{A, B, C, D\} \quad \textit{Variables}$$
$$E = \{(B \to A), (B \to D), (C \to D)\} \quad \textit{Edges}$$
$$P(A, B, C, D) = P(B)P(C)P(A|B)P(D|B, C) \ \textit{Fact.}$$

(b) Bayesian Network




$$V = \{A, B, C, D\} \quad \textit{Variables}$$
$$E = \{(B - A), (B - D), (C - D)\} \quad \textit{Edges}$$
$$P(A, B, C, D) = \frac{1}{Z_\phi} \phi(A, B)\phi(B, D)\phi(C, D) \quad \textit{Fact.}$$
$$Z_\phi = \sum_{A,B,C,D} \phi(A, B)\phi(B, D)\phi(C, D) \quad \textit{Norm.}$$

(c) Markov Random Field




$$V = \{A, B, C, D\} \quad \textit{Variables}$$
$$E = \{(A \leftrightarrow B), (B \to D), (C \to D)\} \quad \textit{Edges}$$
$$f(A, B, C, D) = f(A, B)f(D|B, C) \quad \textit{Factorization}$$

(d) Acyclic Directed Mixed Graph

**Figure 3.1**: Some of the most prominent causal models in literature are Structural Causal Models, Bayesian Neworks, Markov Random Fields and Acyclic Directed Mixed Graphs.

## 4.1 Causal Discovery Methods using Interventional Data

This family of discovery methods uses interventions to estimate the direction between two variables. In fact, the term intervention is a cover for many other definitions. It most often stands for perfect interventions, that only affect the desired events and relations. But the term could also be interpreted as imperfect intervention [31], stochastic intervention [16], soft intervention [17], unreliable intervention [6] or uncertain intervention [6]. When we use the term intervention in this report, we refer to perfect interventions.
The most prominent model for interventional deduction is the Potential Outcomes Framework [28]. It was later found to be the main ingredient of Randomized Controlled Trials and A/B tests, the two most popular methods at the time. The PO Framework comes with the $do$-calculus developed by Judea Pearl [25, 23]. It is a way of denoting, which variables are intervened on. Usually, this method assumes perfect interventions.
To investigate the causal relation between two variables, two randomized sample groups are treated with the potential cause $t_1$ and an alternative $t_2$. Then we compare the outcome of both groups by calculating the causal effect $Y_{t_1} - Y_{t_2}$. Imagine an experiment where we want to find out the effect of a drug $T$. We do this by intervening in the place of the potential remedy $do(T)$. Thereby, $T$ can assume the values 'give medicine' $do(T = t_1)$ or 'give no medicine' $do(T = t_2)$. Then we measure the effects $Y_{t_1}$ and $Y_{t_2}$ in the two independent randomized test groups. For example, the causal effect could be the difference in recovered participants. If the same count of people recover in each test group, the medicine had no effect on the recovery.

## 4.2 Causal Discovery Methods using Observational Data

This group of methods uses observations of the system behavior. This avoids the effort and cost that come with interventions.
Several groups of methods can learn the causal structure in this way: Common are constraint-based, score-based and other functional methods. Each one makes use of certain statistical patterns in the data. The most important patterns are from conditional independencies [22, 8]. The group of score-based discovery methods

calculates a model fit score from the data. In the next step, it optimizes the score by searching in the space of possible models. Which model the presented method finds is shown in Table 4.1. An example is the Iterative Conditional Mode algorithm [2], which maximizes the fit likelihood of a generated model same as Greedy Equivalence Search (GES). Whereby, GES searches directly in the space of Markov equivalence classes [3], a set of models which express the same conditional independencies.
The constraint-based methods use first conditional independence tests and an edge orientation phase to learn the causal model. Popular are the Inductive Causation (IC) algorithm [33, 24] and the Fast Causal Inference (FCI) algorithm [30]. The disadvantage of such methods result from the conditional independence tests. A conditional independence is not always certain and the tests can require a large amount of data to be faithful.
Besides these groups exists a collection of methods exploiting other properties of causal relations. For example, the Additive Noise Model (ANM) makes use of the independence condition as it only holds true in the causal direction [11]. Opposed to the other method groups, that a false learned relation will not effect all other causal relations of the model [8].
In general, causal discovery based on observations alone can rarely discover the whole true causal graph (ancestral graph). Most often, they are limited to the level of the Markov equivalence. One way to overcome this is by means of interventions [32].

## 4.3 Causal Discovery Methods using Observational & Interventional Data

In recent years, methods have become popular, which use data created by interventions and observations. Some methods use them in one joint pool and directly construct their causal model from it. While other methods use each kind of data separately to reconstruct the causal graph.
An example pooling method is Joint Causal Inference (JCI) [22]. It is actually a group of methods, as it joins the data from different contexts, preprocesses it and then allows any other common causal discovery method as IC or FCI to construct the causal graph.

**Table 4.1**: Lists of Observational and Mixed Algorithms that can Discover the Respective Causal Models

| Causal Model | Causal Discovery Methods |
|---|---|
| DAG | ANM [11] |
| | IC [33, 24] |
| SCM | JCI [22] |
| BN | GES [3] |
| | JCI [22] |
| MRF | ICM [2] |
| ADMG | JCI [22] |

Other popular pooling methods include [31, 6, 4, 32]. Splitting methods may take diverse forms. The easiest way is to first use observational methods til the Markov equivalence level and then continue by using interventions.

# 5 Causal Discovery in Technical Systems

In the previous sections, we provided a compact overview over existing methods in causal discovery. Here, we want to give an insight in how we plan on putting them to use in a technical system.

A technical system may take various forms. For example it can be of electrical, water-driven or material-driven nature. Per definition, each artificial system in which matter, power and information interact is a technical system. We assume such a system to be representable in a causal model. From the models described before, ADMGS are the models that capture most domain knowledge.

SCMs and BNs cannot capture bi-directional edges. While MRFs cannot represent the more frequently occurring unidirectional edges [12]. Yet, ADMGs are the least used of the presented models. Further, ADMGs require cost-expensive calculations for causal inference as with each undirected edge, the calculations become more complex [12]. Hence, we advise to first invest in BNs in probabilistic scenarios, and into SCMs or DAGs in non-probabilistic scenarios, before taking the step to ADMGs.

Concerning the discovery methods, we plan to exceed the Markov Equivalence level by using interventions. But as experiments based on the PO Framework examine only a small number of variables [28], it is too expensive to analyze an entire network. Hence, we have to either take observational methods into account or come up with a new method.

# 6 Research Gap and Proposed Approach

Here, we propose a new form of query which considers such indirect causal relations as of $A$ on $C$. The basic idea is that we can observe the spread of a change caused by an intervention. To see which intervention has caused which variable changes should enable us to draw conclusions about the causal structure. By collecting several such constraints and by using combinatorial analysis, we hope to deduce the full causal graph. As a side outcome, we might be able to receive such new information as the dead time between cause and effect variables.
The main effort will be detecting the propagating change in the system. We have identified two options. For one, we could compare the state under intervention with a normal state. If the variable of the intervened state deviates from the corresponding variable of the normal state, the variable indirectly depends on the intervention.
In the second option, the intervention creates a kind of signal, which propagates through the system. By trying to recover this signal from the other variables, we detect which variables depend on the intervened variable. We assume this methods to be more difficult, as the signal is likely to change its form when traveling through the system.
For both options, we have to assume consistency in the system environment as external influences on the system can lead to false conclusions in the causal order.
If this new form of query works out, we could use active learning methods to optimize the costs of intervention and the knowledge gain . Several methods such methods for PO queries exist[32, 9]. We would need a new active learning

method for our query.
As a step further in the future, we see the removal of various constraints we impose on the system. For example, we assume our system to be non-cyclic, but also cyclic graphs have been studied in literature [21, 20]. Also we assume to have no influences on the system besides our interventions. It is likely that in a real system this will not be the case and we will have a greater problem in finding created changes.
For the beginning, we will make use of existing simulations of technical systems as the Tennessee Eastman Process [1, 19]. Such simulations are easy to calculate and allow rapid progress in the development of new algorithms. They are already mathematically reproduced and thus causally inferred, what also allows our models to be easily validated. This allows an easy entry into the development of causal discovery methods for technical systems before we roll them out on real plants.

# 7 Conclusion

In this report, we offered a survey on existing methods in structure learning and causal discovery. If it is convenient, such presented methods could be used for causal inference in technical systems, since there are different possibilities and risks than in domains as the social sciences.
A new approaches for the investigation in observational and interventional causal discovery were proposed. We advice for further investigation to consider indirect causal relations and to use combinatorial inference to learn the causal model structure. To discern the effects of interventions from their environment, we proposed to use comparisons with a most similar twin system or to investigate the injection and tracking of a short signal.

# References

[1] Andreas Bathelt, N. Lawrence Ricker, and Mohieddine Jelali. "Revision of the Tennessee Eastman Process Model". In: *IFAC-PapersOnLine*

48.8 (2015). 9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015, pp. 309–314. DOI: 10.1016/j.ifacol.2015.08.199.

[2] Julian Besag. "On the statistical analysis of dirty pictures". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 48 (1986), pp. 259–279. DOI: 10.1111/j.2517-6161.1986.tb01412.x.

[3] David Maxwell Chickering. "Optimal structure identification with greedy search". In: *Journal of machine learning research* 3 (2002), pp. 507–554.

[4] Gregory F Cooper and Changwon Yoo. "Causal discovery from a mixture of experimental and observational data". In: Morgan Kaufmann Publishers Inc., 1999. DOI: 10.5555/2073796.2073810.

[5] Paul Dagum, Adam Galper, and Eric Horvitz. "Dynamic Network Models for Forecasting". In: *Uncertainty in Artificial Intelligence*. Ed. by Didier Dubois et al. Morgan Kaufmann, 1992, pp. 41–48. ISBN: 978-1-4832-8287-9. DOI: 10.1016/B978-1-4832-8287-9.50010-4.

[6] Daniel Eaton and Kevin Murphy. "Exact Bayesian structure learning from uncertain interventions". In: *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*. Ed. by Marina Meila and Xiaotong Shen. Vol. 2. Proceedings of Machine Learning Research. PMLR, 2007, pp. 107–114. URL: http://proceedings.mlr.press/v2/eaton07a.html.

[7] Clark Glymour, Kun Zhang, and Peter Spirtes. "Review of causal discovery methods based on graphical models". In: vol. 10. Frontiers, 2019, p. 524.

[8] Olivier Goudet et al. "Learning functional causal models with generative neural networks". In: *Explainable and Interpretable Models in Computer Vision and Machine Learning*. Springer, 2018, pp. 39–80.

[9] Yang-Bo He and Zhi Geng. "Active Learning of Causal Networks with Intervention Experiments and Optimal Designs". In: *Journal of Machine Learning Research* 9 (2008), pp. 2523–2547. URL: http://jmlr.org/papers/v9/he08a.html.

[10] Paul W Holland. "Statistics and causal inference". In: vol. 81. Taylor & Francis, 1986, pp. 945–960.

[11] Patrik Hoyer et al. "Nonlinear causal discovery with additive noise models". In: *Advances in Neural Information Processing Systems*. Ed. by D. Koller et al. Vol. 21. Curran Associates, Inc., 2009, pp. 689–696. URL: https://proceedings.neurips.cc/paper/2008/file/f7664060cc52bc6f3d620bcedc94a4b6-Paper.pdf.

[12] Pearl Judea. "Causality: models, reasoning, and inference". In: 521 (2000), p. 8.

[13] Markus Kalisch and Peter Bühlmann. "Causal structure learning and inference: a selective review". In: *Journal of Economic Methodology* 11 (2014), pp. 81–98. DOI: 10.1080/16843703.2014.11673322.

[14] Ross Kindermann. "Markov random fields and their applications". In: *American mathematical society* (1980). DOI: 10.1090/conm/001.

[15] Daphne Koller and Avi Pfeffer. "Object-oriented Bayesian networks". In: *Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence* (1997), pp. 302–313.

[16] Kevin Korb et al. "Varieties of causal intervention". In: ed. by Yeap WK. Zhang C. W. Guesgen H. Vol. 3157. Springer. 2004, pp. 322–331. DOI: 10.1007/978-3-540-28633-2_35.

[17] Christian Kuehnert, Thomas Bernard, and Christian Frey. "Causal structure learning in process engineering using Bayes Nets and soft interventions". In: *2011 9th IEEE International Conference on Industrial Informatics*. IEEE. 2011. DOI: 10.1109/INDIN.2011.6034839.

[18] Hans Kunsch, Stuart Geman, and Athanasios Kehagias. "Hidden Markov Random Fields". In: *Annals of Applied Probability* 5.3 (Aug. 1995), pp. 577–602. DOI: 10.1214/aoap/1177004696.

[19] Carla Martin-Villalba, Alfonso Urquia, and Guodong Shao. "Implementations of the Tennessee Eastman Process in Modelica". In: *IFAC-PapersOnLine* 51 (2018). 9th Vienna International Conference on Mathematical Modelling, pp. 619–624. ISSN: 2405-8963. DOI: 10.1016/j.ifacol.2018.03.105.

[20] Joris M Mooij et al. “On Causal Discovery with Cyclic Additive Noise Models”. In: *Advances in Neural Information Processing Systems*. Ed. by J. Shawe-Taylor et al. Vol. 24. Curran Associates, Inc., 2011, pp. 639–647. URL: https://proceedings.neurips.cc/paper/2011/file/d61e4bbd6393c9111e6526ea173a7c8b-Paper.pdf.

[21] Joris Mooij and Tom Heskes. “Cyclic causal discovery from continuous equilibrium data”. In: *arXiv:1309.6849* (2013).

[22] Joris Mooij, Sara Magliacane, and Tom Claassen. “Joint causal inference from multiple contexts”. In: *arXiv preprint arXiv:1611.10351* (2016).

[23] Judea Pearl. “Causal diagrams for empirical research”. In: *Biometrika* 82 (1995).

[24] Judea Pearl. *Causality*. Cambridge university press, 2009. DOI: 10.1017/CBO9780511803161.

[25] Judea Pearl. “The do-calculus revisited”. In: *arXiv:1210.4852* (2012).

[26] Judea Pearl, Madelyn Glymour, and Nicholas P Jewell. *Causal inference in statistics: A primer*. John Wiley & Sons, 2016.

[27] Jose Peña. “Alternative Markov and causal properties for acyclic directed mixed graphs”. In: *arXiv preprint arXiv:1511.05835* (2015). DOI: 10.1023/A:1017445827962.

[28] Donald B Rubin. “Causal inference using potential outcomes: Design, modeling, decisions”. In: *Journal of the American Statistical Association* 100 (2005), pp. 322–331. DOI: 10.1198/016214504000001880.

[29] Havard Rue and Leonhard Held. *Gaussian Markov random fields: theory and applications*. Chapman and Hall, 2005. DOI: 10.1201/9780203492024.

[30] Pater Spirtes et al. “Constructing Bayesian network models of gene expression networks from microarray data”. In: (2000).

[31] Jin Tian and Judea Pearl. “Causal Discovery from Changes”. In: (2001), pp. 512–521. DOI: 10.5555/2074022.2074085.

[32] Simon Tong and Daphne Koller. “Active learning for structure in Bayesian networks”. In: *International joint conference on artificial intelligence*. Vol. 17. 2001, pp. 863–869.

[33] Thomas Verma and Judea Pearl. *Equivalence and synthesis of causal models*. UCLA, Computer Science Department, 1991.