# Identifying Dynamic Sequential Plans

Jin Tian
jtian@cs.iastate.edu
(Presented by George Voutsadakis)

Department of Computer Science
Iowa State University, Ames, IA 50011, USA

The 24th Conference on Uncertainty in Artificial Intelligence

1. Background
   - Bayesian Networks
   - Causal Bayesian Networks
   - Atomic Interventions and Causal Effects
2. Sequential Decision Plans and Intervention Strategies
   - Sequential Decision Plans
   - Intervention Strategies; The Intervention Graph
3. The Identifiability Problem
   - Definition of Identifiability
   - The Main Theorem
4. Conclusions

# Bayesian Networks

A **Bayesian network** $G = \langle W, E \rangle$ is a directed graph.

- Nodes represent variables;
- Edges represent conditional independences: each variable is independent of all its non-descendants given its direct parents in $G$.

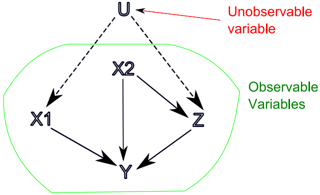



Figure: A Bayesian Network.

$$P(y, z, x_1, x_2, u) = P(y|x_1, x_2, z)P(z|x_2, u)P(x_1|u)P(x_2)P(u).$$

$$P(y, z, x_1, x_2) = \sum_u P(y|x_1, x_2, z)P(z|x_2, u)P(x_1|u)P(x_2)P(u).$$

# Causal Bayesian Networks

A **causal Bayesian network** $G$ is a Bayesian network, such that, in equation

$$P^M(w) = \prod_{W_i \in W} P^M(w_i | \mathrm{pa}_{w_i})$$

# Causal Bayesian Networks

A **causal Bayesian network** $G$ is a Bayesian network, such that, in equation

$$P^M(w) = \prod_{W_i \in W} P^M(w_i | \mathrm{pa}_{w_i})$$

1. **Causality**: The stochastic variation of the assignment $P^M(w_i|\mathrm{pa}_{w_i})$ is chosen in response to the values $\mathrm{pa}_{w_i}$ and is independent of the variations in all other assignments;

# Causal Bayesian Networks

A **causal Bayesian network** $G$ is a Bayesian network, such that, in equation

$$P^M(w) = \prod_{W_i \in W} P^M(w_i | \mathrm{pa}_{w_i})$$

1. **Causality**: The stochastic variation of the assignment $P^M(w_i|\mathrm{pa}_{w_i})$ is chosen in response to the values $\mathrm{pa}_{w_i}$ and is independent of the variations in all other assignments;
2. **Modularity**: Each assignment remains invariant under possible changes in the assignments governing other variables in the system.

# Atomic Interventions

- **Atomic intervention** $\mathrm{do}(X_i = x_i^*)$ to a variable $X_i$ in a Bayesian network: For some fixed value $x_i^*$,

$$P(x_i|\mathrm{pa}_{x_i}) = \delta(x_i, x_i^*) := \begin{cases} 1, & \text{if } x_i = x_i^* \\ 0, & \text{otherwise} \end{cases} .$$

# Atomic Interventions

- **Atomic intervention** $\mathrm{do}(X_i = x_i^*)$ to a variable $X_i$ in a Bayesian network: For some fixed value $x_i^*$,

$$P(x_i|\mathrm{pa}_{x_i}) = \delta(x_i, x_i^*) := \begin{cases} 1, & \text{if } x_i = x_i^* \\ 0, & \text{otherwise} \end{cases} .$$

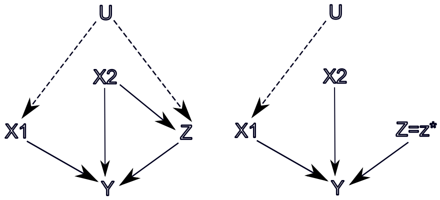



Figure: Atomic Intervention on $Z$.

# Causal Effects

- The **causal effect** $P_x(y)$ that an intervention on $X$ has on the outcome variables $Y$ can be computed by

$$P_x(y) = \begin{cases} \sum_{z,u} \prod_{Y_i \in Y} P(y_i|\mathrm{pa}_{y_i}) \prod_{Z_i \in Z} P(z_i|\mathrm{pa}_{z_i}) \prod_{U_i \in U} P(u_i|\mathrm{pa}_{u_i}), & \\ \qquad \text{if } y \text{ consistent with } x & \\ 0, & \text{otherwise} \end{cases}$$

# Causal Effects

- The **causal effect** $P_x(y)$ that an intervention on $X$ has on the outcome variables $Y$ can be computed by

$$P_x(y) = \begin{cases} \sum_{z,u} \prod_{Y_i \in Y} P(y_i|\mathrm{pa}_{y_i}) \prod_{Z_i \in Z} P(z_i|\mathrm{pa}_{z_i}) \prod_{U_i \in U} P(u_i|\mathrm{pa}_{u_i}), & \text{if } y \text{ consistent with } x \\ 0, & \text{otherwise} \end{cases}$$

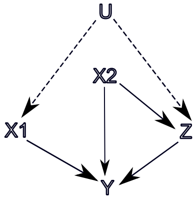



Figure: Atomic Interventions and Causal Effects.

$$P_{x_1^*, x_2^*}(y) = \sum_{z,u} P(y|x_1^*, x_2^*, z) P(z|x_2^*, u) P(u).$$

# Sequential Decision Plans

A **sequential decision plan** $S = \langle G, \langle X, Y, Z \rangle \rangle$ consists of

# Sequential Decision Plans

A **sequential decision plan** $S = \langle G, \langle X, Y, Z \rangle \rangle$ consists of

- a Bayesian network with unobservable variables $G = \langle V \cup U, E \rangle$;

# Sequential Decision Plans

A **sequential decision plan** $S = \langle G, \langle X, Y, Z \rangle \rangle$ consists of

- a Bayesian network with unobservable variables $G = \langle V \cup U, E \rangle$;
- A partition $V = X \cup Y \cup Z$ of its set of observable variables into
  - a set $X$ of **action variables**,
  - a set $Y$ of **outcome variables** and
  - a set $Z$ of **covariate variables**;

# Sequential Decision Plans

A **sequential decision plan** $S = \langle G, \langle X, Y, Z \rangle \rangle$ consists of

- a Bayesian network with unobservable variables $G = \langle V \cup U, E \rangle$;
- A partition $V = X \cup Y \cup Z$ of its set of observable variables into
  - a set $X$ of **action variables**,
  - a set $Y$ of **outcome variables** and
  - a set $Z$ of **covariate variables**;
- An intervention strategy $\sigma_X = \{\sigma_{X_i}\}_{X_i \in X}$.

# Sequential Decision Plans

A **sequential decision plan** $S = \langle G, \langle X, Y, Z \rangle\rangle$ consists of

- a Bayesian network with unobservable variables $G = \langle V \cup U, E \rangle$;
- A partition $V = X \cup Y \cup Z$ of its set of observable variables into
  - a set $X$ of **action variables**,
  - a set $Y$ of **outcome variables** and
  - a set $Z$ of **covariate variables**;
- An intervention strategy $\sigma_X = \{\sigma_{X_i}\}_{X_i \in X}$.

Several kinds of intervention strategies may be applied to the action variables.

# Possible Intervention Strategies for the Action Variables

Each intervention to the action variables may be of one of the following types:

- **Idle Intervention:** $P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i}) = P(x_i|\mathrm{pa}_{x_i})$.

# Possible Intervention Strategies for the Action Variables

Each intervention to the action variables may be of one of the following types:

- **Idle Intervention:** $P(x_i|\mathrm{pa}_{x_i};\sigma_{X_i}) = P(x_i|\mathrm{pa}_{x_i})$.
- **Atomic Intervention:** $P(x_i|\mathrm{pa}_{x_i};\sigma_{X_i} = \mathrm{do}(x_i^*)) = \delta(x_i, x_i^*)$

# Possible Intervention Strategies for the Action Variables

Each intervention to the action variables may be of one of the following types:

- **Idle Intervention:** $P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i}) = P(x_i|\mathrm{pa}_{x_i})$.
- **Atomic Intervention:** $P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i} = \mathrm{do}(x_i^*)) = \delta(x_i, x_i^*)$
- **Conditional Intervention:** $P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i} = \mathrm{do}(g(c))) = \delta(x_i, g(c))$, where $g(\cdot)$ is a prespecified deterministic function and the variables in $C$ cannot be descendants of $X_i$.

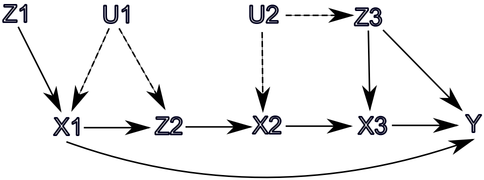
# Possible Intervention Strategies for the Action Variables

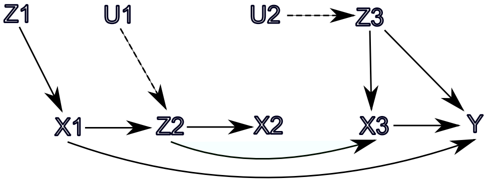
Each intervention to the action variables may be of one of the following types:

- **Idle Intervention:** $P(x_i|\mathrm{pa}_{x_i};\sigma_{X_i}) = P(x_i|\mathrm{pa}_{x_i})$.
- **Atomic Intervention:** $P(x_i|\mathrm{pa}_{x_i};\sigma_{X_i} = \mathrm{do}(x_i^*)) = \delta(x_i, x_i^*)$
- **Conditional Intervention:** $P(x_i|\mathrm{pa}_{x_i};\sigma_{X_i} = \mathrm{do}(g(c))) = \delta(x_i, g(c))$, where $g(\cdot)$ is a prespecified deterministic function and the variables in $C$ cannot be descendants of $X_i$.
- **Random Intervention:** $P(x_i|\mathrm{pa}_{x_i};\sigma_{X_i} = d_C) = P^*(x_i|c)$, where $P^*(x_i|c)$ is a prespecified probability distribution and the variables in $C$ cannot be descendants of $X_i$.

# The Intervention Graph $G_{\sigma_X}$

Consider the top graph $G$ below. Suppose that we follow the intervention strategies: $\sigma_{X_1} = \mathrm{do}(g_1(Z_1)), \sigma_{X_2} = d_{Z_2}, \sigma_{X_3} = \mathrm{do}(g_3(Z_2, Z_3))$.




The Intervention graph $G_{\sigma_X}$ corresponding to $\sigma_X = \{\sigma_{X_1}, \sigma_{X_2}, \sigma_{X_3}\}$ is the graph given at the bottom above.

# Example:

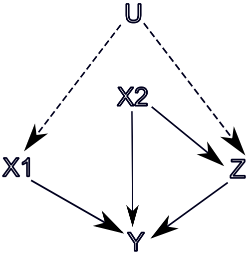



Figure: Sequential Decision Plans.

$$P(v) = \sum_{u} P(y|x_1, x_2, z)P(z|x_2, u)P(x_1|u)P(x_2)P(u).$$

$$P(v; \sigma_X) = \sum_{u} P(y|x_1, x_2, z)P(z|x_2, u)P(x_1|u; \sigma_{X_1})P(x_2; \sigma_{X_2})P(u).$$

$$P(y; \sigma_X) = \sum_{x_1, x_2, z, u} P(y|x_1, x_2, z)P(z|x_2, u)P(x_1|u; \sigma_{X_1})P(x_2; \sigma_{X_2})P(u).$$

# Definition of Identifiability

- **Identifiability** means that the join probability distribution of the observable variables (without interventions) uniquely determines the post-interventional join probability distribution of the outcome variables:

$$P^{M_1}(v) = P^{M_2}(v) \quad \text{implies} \quad P^{M_1}(y; \sigma_X) = P^{M_2}(y; \sigma_X).$$

# Definition of Identifiability

- **Identifiability** means that the join probability distribution of the observable variables (without interventions) uniquely determines the post-interventional join probability distribution of the outcome variables:

$$P^{M_1}(v) = P^{M_2}(v) \quad \text{implies} \quad P^{M_1}(y; \sigma_X) = P^{M_2}(y; \sigma_X).$$

**The Identifiability Problem**

Given a sequential decision plan, determine whether it is identifiable.

# Definition of Identifiability

- **Identifiability** means that the join probability distribution of the observable variables (without interventions) uniquely determines the post-interventional join probability distribution of the outcome variables:

$$P^{M_1}(v) = P^{M_2}(v) \quad \text{implies} \quad P^{M_1}(y; \sigma_X) = P^{M_2}(y; \sigma_X).$$

**The Identifiability Problem**

Given a sequential decision plan, determine whether it is identifiable.

- Identifiability means that $P(y; \sigma_X)$ is computable from the causal graph (which may contain unobservable variable) and the join distribution $P(v)$ of the observable variables.

# Known Solution:

## Causal Effects

The identifiability problem has been completely solved for the case of sequential decision plans with all interventions being atomic, i.e., the case of *causal effects*. (Tian and Pearl, Shpitser and Pearl, Huang and Valtorta)

# Known Solution:

## Causal Effects

The identifiability problem has been completely solved for the case of sequential decision plans with all interventions being atomic, i.e., the case of *causal effects*. (Tian and Pearl, Shpitser and Pearl, Huang and Valtorta)

## The General Case

The general case is reduced in this work to the identifiability of causal effects. Thus the identifiability of arbitrary sequential decision plans is resolved.

# The Main Theorem

## Key Assumption:

For a conditional intervention $\sigma_{X_i} = \mathrm{do}(g(c))$ or a random intervention $\sigma_{X_i} = d_C$, it is required that $C \subseteq X \cup Z$, i.e., for every $X_i \in X$, $P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i})$ will not depend on the unobservable variables.

# The Main Theorem

## Key Assumption:

For a conditional intervention $\sigma_{X_i} = \mathrm{do}(g(c))$ or a random intervention $\sigma_{X_i} = d_C$, it is required that $C \subseteq X \cup Z$, i.e., for every $X_i \in X$, $P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i})$ will not depend on the unobservable variables.

## Theorem

*Let $S = \langle G, \langle X, Y, Z \rangle \rangle$ be a sequential decision plan, $\sigma_X = \{\sigma_{X_i}\}$ an intervention strategy and denote by $X_D, Z_D$ the subsets of $X$ and $Z$, respectively, that are ancestors of $Y$ in $G_{\sigma_X}$. If the causal effect $P_{x,z \setminus z_D}(y, z_D)$ is identifiable, then $P(y; \sigma_X)$ is identifiable using the formula*

$$P(y; \sigma_X) = \sum_{x_D, z_D} \prod_{i: X_i \in X_D} P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i}) P_{x,z \setminus z_D}(y, z_D).$$

# The Main Theorem

## Key Assumption:

For a conditional intervention $\sigma_{X_i} = \mathrm{do}(g(c))$ or a random intervention $\sigma_{X_i} = d_C$, it is required that $C \subseteq X \cup Z$, i.e., for every $X_i \in X$, $P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i})$ will not depend on the unobservable variables.

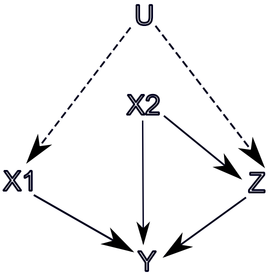
## Theorem

*Let $S = \langle G, \langle X, Y, Z \rangle \rangle$ be a sequential decision plan, $\sigma_X = \{\sigma_{X_i}\}$ an intervention strategy and denote by $X_D, Z_D$ the subsets of $X$ and $Z$, respectively, that are ancestors of $Y$ in $G_{\sigma_X}$. If the causal effect $P_{x, z \setminus z_D}(y, z_D)$ is identifiable, then $P(y; \sigma_X)$ is identifiable using the formula*

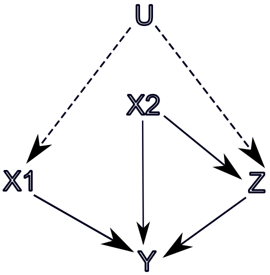
$$P(y; \sigma_X) = \sum_{x_D, z_D} \prod_{i: X_i \in X_D} P(x_i|\mathrm{pa}_{x_i}; \sigma_{X_i}) P_{x, z \setminus z_D}(y, z_D).$$

How about the Converse?

# Example (Dawid and Didelez)

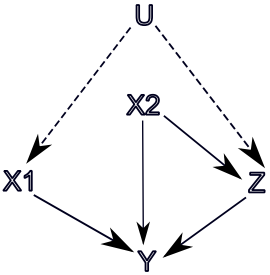


# Example (Dawid and Didelez)




- $P(y;\sigma_X) = \sum_{x_1,x_2,z} P(x_1;\sigma_{X_1})P(x_2;\sigma_{X_2})P_{x_1,x_2}(y,z)$.

# Example (Dawid and Didelez)




- $P(y;\sigma_X) = \sum_{x_1,x_2,z} P(x_1;\sigma_{X_1})P(x_2;\sigma_{X_2})P_{x_1,x_2}(y,z)$.
- $P_{x_1,x_2}(y,z)$ may be shown to be identifiable

$$P_{x_1,x_2}(y,z) = P(y|x_1,x_2,z)P(z,x_1|x_2)$$

and, therefore $P(y;\sigma_X)$ is also identifiable.

# Conclusions

- A method is presented for identifying dynamic sequential plans based on the identifiability of causal effects.
- Since algorithms for the identifiability of causal effects are available in the literature, this method is useful in practice.