Causal Discovery

Adapted from CAUSALITY[1] chapter 2

Ofek Gila[1]

[1]University of California, Irvine

February 2022

Table of Contents

Motivation

I would rather discover one cause than gain the kingdom of Persia.
—Democritus (460-370 B.C.)[2]

Motivation

I would rather discover one cause than gain the kingdom of Persia.
—Democritus (460-370 B.C.)[2]

- Heavily sought after in 1700s

Motivation

I would rather discover one cause than gain the kingdom of Persia.
—Democritus (460-370 B.C.)[2]

- Heavily sought after in 1700s
- Feasibly computable in mid-1980s

Motivation

I would rather discover one cause than gain the kingdom of Persia.
—Democritus (460-370 B.C.)[2]

- Heavily sought after in 1700s
- Feasibly computable in mid-1980s

Why take so long?

Motivation

I would rather discover one cause than gain the kingdom of Persia.
—Democritus (460-370 B.C.)[2]

- Heavily sought after in 1700s
- Feasibly computable in mid-1980s

Why take so long?

Statistical associations $\not\to$ Causation

Preface

- Statistical associations $\not\rightarrow$ Causation

Preface

- Statistical associations $\not\rightarrow$ Causation
- Statistical associations *suggest* Causation?

Preface

- Statistical associations $\not\rightarrow$ Causation
- Statistical associations *suggest* Causation?

1. What clues allow people perceive causal relationships from uncontrolled observations?

Preface

- Statistical associations $\not\rightarrow$ Causation
- Statistical associations *suggest* Causation?

1. What clues allow people perceive causal relationships from uncontrolled observations?
2. What assumptions would allow us to infer causal models from these clues?

Preface

- Statistical associations $\not\rightarrow$ Causation
- Statistical associations *suggest* Causation?

1. What clues allow people perceive causal relationships from uncontrolled observations?
2. What assumptions would allow us to infer causal models from these clues?
3. Will the inferred models be useful?

Table of Contents

Intuition

- Autonomous intelligent system
  - Requires workable environment model

Intuition

- Autonomous intelligent system
  - Requires workable environment model
- Cannot be preprogrammed w/ all causal knowledge

Intuition

- Autonomous intelligent system
  - Requires workable environment model
- Cannot be preprogrammed w/ all causal knowledge
- Must rely on direct observations

Intuition

- Autonomous intelligent system
  - Requires workable environment model
- Cannot be preprogrammed w/ all causal knowledge
- Must rely on direct observations
- Is this possible?

Intuition

- Autonomous intelligent system
  - Requires workable environment model
- Cannot be preprogrammed w/ all causal knowledge
- Must rely on direct observations
- Is this possible?

No?

Covariation $\not\rightarrow$ Causation

Intuition

- Autonomous intelligent system
  - Requires workable environment model
- Cannot be preprogrammed w/ all causal knowledge
- Must rely on direct observations
- Is this possible?

No?

Covariation $\not\rightarrow$ Causation

Yes?

But humans do it!

Temporal Knowledge

- Temporal precedence?

Temporal Knowledge

- Temporal precedence?
- $t_{\text{cause}} < t_{\text{effect}}$

Temporal Knowledge

- Temporal precedence?
- $t_{\text{cause}} < t_{\text{effect}}$

Thought experiment

Let A and B be dependent. Let $t_A < t_B$. Does $A \rightarrow B$?

Temporal Knowledge

- Temporal precedence?
- $t_{\text{cause}} < t_{\text{effect}}$

Thought experiment

Let A and B be dependent. Let $t_A < t_B$. Does $A \to B$?

No!

- A is rooster crowing, B is sun going up
- A is barometer falling, B is it raining outside

Then What?

- Patterns indicative of causal directionality?

Then What?

- Patterns indicative of causal directionality?

Another thought experiment

Let A and B be dependent, B and C be dependent, yet A and C be independent. What is the SCM?

Then What?

- Patterns indicative of causal directionality?

Another thought experiment

Let A and B be dependent, B and C be dependent, yet A and C be independent. What is the SCM?

A and B independently cause C, i.e., $A \rightarrow B \leftarrow C$

Then What?

- Patterns indicative of causal directionality?

Another thought experiment

Let A and B be dependent, B and C be dependent, yet A and C be independent. What is the SCM?

A and B independently cause C, i.e., $A \rightarrow B \leftarrow C$

Can B cause A and C?

Then What?

- Patterns indicative of causal directionality?

Another thought experiment

Let A and B be dependent, B and C be dependent, yet A and C be independent. What is the SCM?

A and B independently cause C, i.e., $A \rightarrow B \leftarrow C$

Can B cause A and C?

Yes but...

Then What?

- Patterns indicative of causal directionality?

Another thought experiment

Let A and B be dependent, B and C be dependent, yet A and C be independent. What is the SCM?

A and B independently cause C, i.e., $A \rightarrow B \leftarrow C$

Can B cause A and C?

Yes but...

- Some patterns *suggest* causality

Framework

Definition

A causal structure of a set of variables V is a directed acyclic graph in which each node corresponds to a distinct element in V, and each link represents a direct functional relationship among the corresponding variables.

Framework

Definition

A causal structure of a set of variables V is a directed acyclic graph in which each node corresponds to a distinct element in V, and each link represents a direct functional relationship among the corresponding variables.

Definition

A causal model is a pair $M = \langle D, \Theta_D \rangle$ consisting of a causal structure D and a set of parameters Θ_D compatible with D. The parameters Θ_D assign a function $x_i = f_i(pa_i, u_i)$ for each X_i in D and where U_i is a random disturbance distributed according to $P(u_i)$ independently of all other u.

Model Preference I

- Unbounded number of models

Model Preference I

- Unbounded number of models
- How to choose?

Model Preference I

- Unbounded number of models
- How to choose? Occam's razor

Model Preference I

- Unbounded number of models
- How to choose? Occam's razor
- Pick the simplest model

Model Preference I

- Unbounded number of models
- How to choose? Occam's razor
- Pick the simplest model

Definition

A latent structure is a pair $L = \langle D, O \rangle$, where D is a causal structure over V and where $O \subseteq V$ is a set of observed variables.

Model Preference I

- Unbounded number of models
- How to choose? Occam's razor
- Pick the simplest model

Definition

A latent structure is a pair $L = \langle D, O \rangle$, where D is a causal structure over V and where $O \subseteq V$ is a set of observed variables.

Definition

One latent structure $L = \langle D, O \rangle$ is preferred to another $L' = \langle D', O \rangle$ ($L \preceq L'$) if and only if D' can mimic D over O. Two latent structures are equal ($L \equiv L'$) if and only if $L \preceq L'$ and $L \succeq L'$

Model Preference II

Definition

One latent structure $L = \langle D, O \rangle$ is preferred to another $L' = \langle D', O \rangle$ ($L \preceq L'$) if and only if D' can mimic D over O. Two latent structures are equal ($L \equiv L'$) if and only if $L \preceq L'$ and $L \succeq L'$

Model Preference II

Definition

One latent structure $L = \langle D, O \rangle$ is preferred to another $L' = \langle D', O \rangle$ ($L \preceq L'$) if and only if D' can mimic D over O. Two latent structures are equal ($L \equiv L'$) if and only if $L \preceq L'$ and $L \succeq L'$

Definition

A latent structure L is minimal with respect to a class $\mathcal{L}$ of latent structures if and only if there is no member of $\mathcal{L}$ that is strictly preferred to L.

Model Preference II

Definition

One latent structure $L = \langle D, O \rangle$ is preferred to another $L' = \langle D', O \rangle$ ($L \preceq L'$) if and only if D' can mimic D over O. Two latent structures are equal ($L \equiv L'$) if and only if $L \preceq L'$ and $L \succeq L'$

Definition

A latent structure L is minimal with respect to a class $\mathcal{L}$ of latent structures if and only if there is no member of $\mathcal{L}$ that is strictly preferred to L.

Definition

A variable A has a causal influence on variable B if and only if there exists a directed path from A to B om every consistent minimal latent structure.

Example I

Example

Given the following observations over variables $\{A, B, C, D\}$:

- A is independent of B
- D is independent of $\{A, B\}$ given C
- There is no other independence

What is the minimal causal structure?

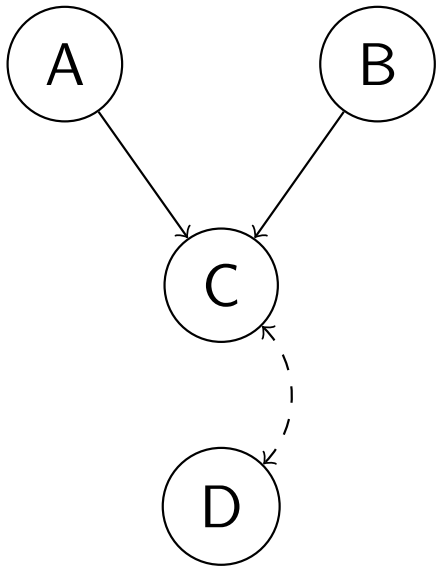
Example I

Example

Given the following observations over variables $\{A, B, C, D\}$:

- A is independent of B
- D is independent of $\{A, B\}$ given C
- There is no other independence

What is the minimal causal structure?



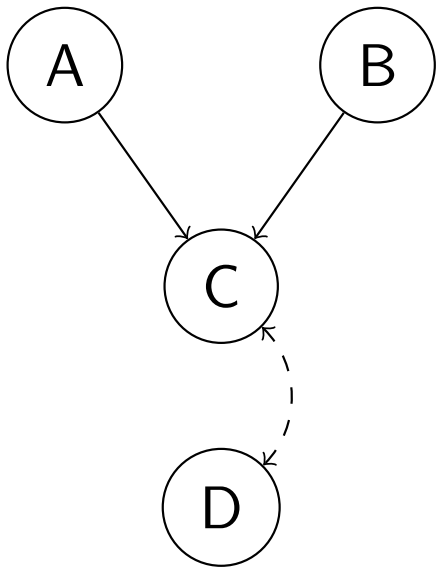
Example I

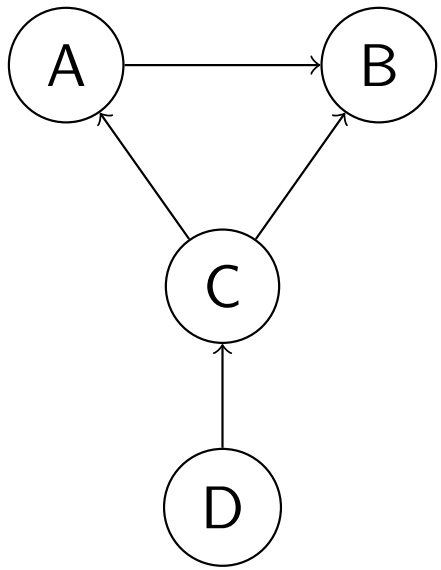
Example

Given the following observations over variables $\{A, B, C, D\}$:

- A is independent of B
- D is independent of $\{A, B\}$ given C
- There is no other independence

What is the minimal causal structure?





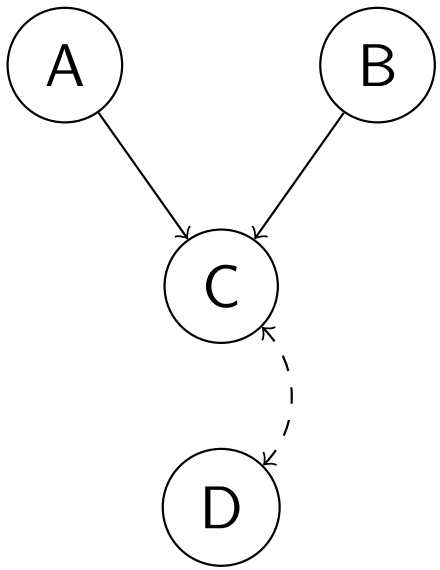
Example I

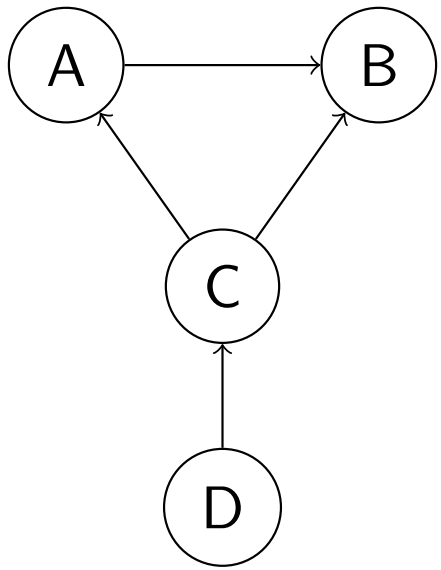
Example

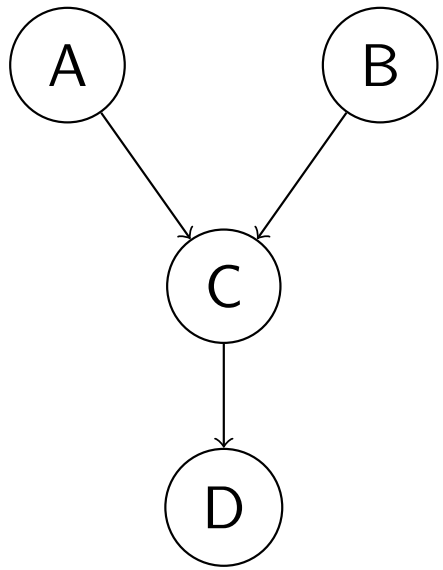
Given the following observations over variables $\{A, B, C, D\}$:

- A is independent of B
- D is independent of $\{A, B\}$ given C
- There is no other independence

What is the minimal causal structure?





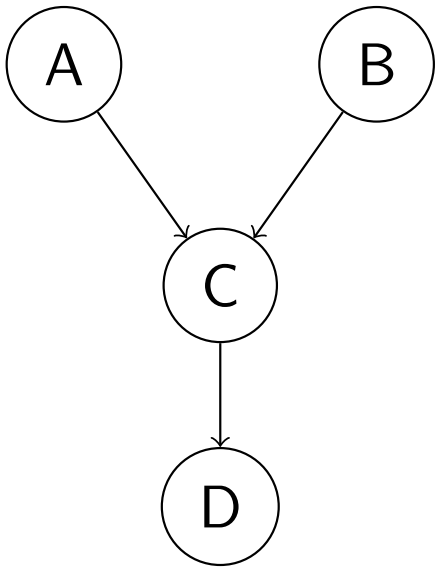


Example II

Example

- A is independent of B
- D is independent of $\{A, B\}$ given C
- There is no other independence

- Unique minimal structure?

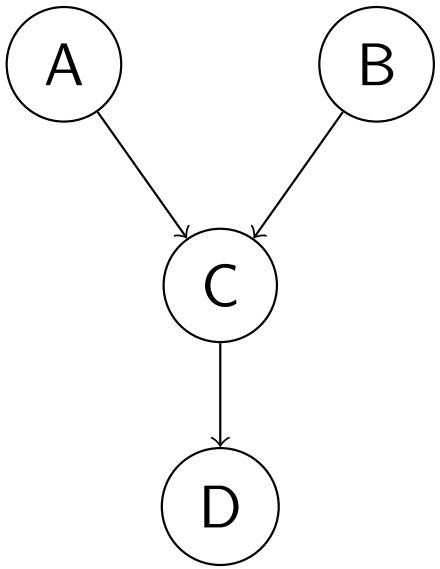


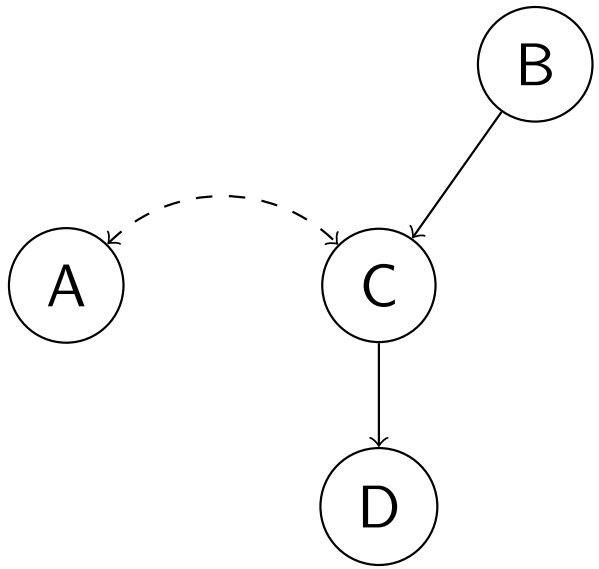
Example II

Example

- A is independent of B
- D is independent of $\{A, B\}$ given C
- There is no other independence

- Unique minimal structure? No





Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?

Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other

Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal

Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

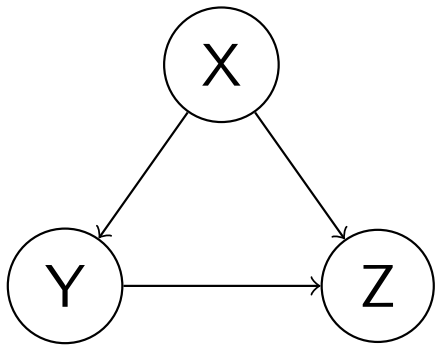
- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal
- Option 2 is unstable to viewing angle, environmental conditions, etc

Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal
- Option 2 is unstable to viewing angle, environmental conditions, etc



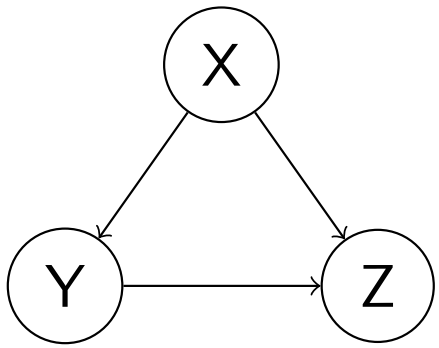
Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal
- Option 2 is unstable to viewing angle, environmental conditions, etc

- What variables are independent?



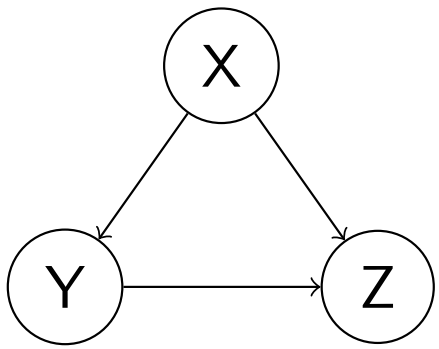
Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal
- Option 2 is unstable to viewing angle, environmental conditions, etc

- What variables are independent?
- Let: $z = \gamma x + u_1, \;\; y = \alpha x + \beta z + u_2$



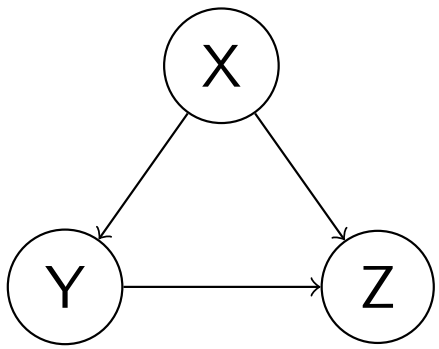
Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal
- Option 2 is unstable to viewing angle, environmental conditions, etc

- What variables are independent?
- Let: $z = \gamma x + u_1, \;\; y = \alpha x + \beta z + u_2$
- What if $\alpha = -\beta\gamma$?



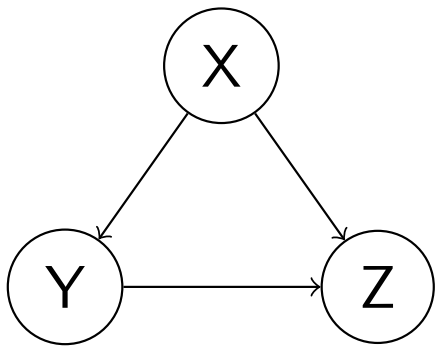
Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal
- Option 2 is unstable to viewing angle, environmental conditions, etc

- What variables are independent?
- Let: $z = \gamma x + u_1, \;\; y = \alpha x + \beta z + u_2$
- What if $\alpha = -\beta\gamma$?
- X and Y are independent?



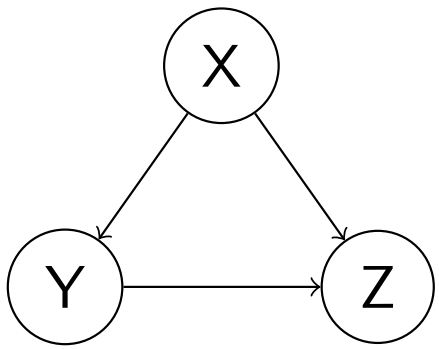
Stability

Definition

A causal model $M = \langle D, \Theta_D \rangle$ is stable if and only if no independence can be destroyed when varying Θ to a Θ'.

- What does this mean?
- Suppose you see a stool in a photo. Is it:
  1. A stool
  2. Two stools positioned so that one hides the other
- Option 1 is a subset of Option 2, minimal
- Option 2 is unstable to viewing angle, environmental conditions, etc

- What variables are independent?
- Let: $z = \gamma x + u_1, \;\; y = \alpha x + \beta z + u_2$
- What if $\alpha = -\beta\gamma$?
- X and Y are independent? unstable



Homework Problem

- Recall:

Another thought experiment

Let A and B be dependent, B and C be dependent, yet A and C be independent. What is the SCM?

A and B independently cause C, i.e., $A \rightarrow B \leftarrow C$

Can B cause A and C?

Yes but...

Homework Problem

- Recall:

Another thought experiment

Let A and B be dependent, B and C be dependent, yet A and C be independent. What is the SCM?

A and B independently cause C, i.e., $A \rightarrow B \leftarrow C$

Can B cause A and C?

Yes but...

Homework

Provide a structural causal model with the above dependencies.
Describe why this model would not be assumed to be the structural causal model using the minimality and stability criteria.

Table of Contents

References I


J. Pearl, *Causality*.
Cambridge: Cambridge University Press, 2 ed., 2009.


K. Freeman, *Ancilla to the Pre-Socratic Philosophers: A Complete Translation of the Fragments in Diels Fragmente Der Vorsokratiker*.
Forgotten Books, 1948.
Google-Books-ID: ASijqFryr5IC.