

On patterns of conjunctive forks

Vašek Chvátal¹, František Matúš and Yori Zwols²

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
matus@utia.cas.cz

Algebraic Statistics

June 8-11, 2015 Genova, Italy

¹Concordia University, Montreal; ²Google, London

H. Reichenbach (1891–1953)

H. Reichenbach (1891–1953)

Stanford Encyclopedia of Philosophy:

“perhaps the greatest empiricist of the 20th century”

H. Reichenbach (1891–1953)

Stanford Encyclopedia of Philosophy:

“perhaps the greatest empiricist of the 20th century”

<http://plato.stanford.edu/entries/reichenbach/>

H. Reichenbach (1891–1953)

Stanford Encyclopedia of Philosophy:

“perhaps the greatest empiricist of the 20th century”

<http://plato.stanford.edu/entries/reichenbach/>

Principle of the Common Cause: “If an improbable coincidence has occurred, there must exist a common cause.”

H. Reichenbach (1891–1953)

Stanford Encyclopedia of Philosophy:

“perhaps the greatest empiricist of the 20th century”

<http://plato.stanford.edu/entries/reichenbach/>

Principle of the Common Cause: “If an improbable coincidence has occurred, there must exist a common cause.”

The Direction of Time (1956) University of California Press.

H. Reichenbach (1891–1953)

Stanford Encyclopedia of Philosophy:

“perhaps the greatest empiricist of the 20th century”

<http://plato.stanford.edu/entries/reichenbach/>

Principle of the Common Cause: “If an improbable coincidence has occurred, there must exist a common cause.”

The Direction of Time (1956) University of California Press.

conjunctive forks play a central role
in Reichenbach’s causal theory of time

An ordered triple (A, B, C) of events in a probability space (Ω, \mathbb{P})

An ordered triple (A, B, C) of events in a probability space (Ω, P) is a **conjunctive fork** if

$$\begin{aligned}P(A \cap C|B) &= P(A|B) \cdot P(C|B), \\P(A \cap C|\Omega \setminus B) &= P(A|\Omega \setminus B) \cdot P(C|\Omega \setminus B), \\P(A|B) &> P(A|\Omega \setminus B), \\P(C|B) &> P(C|\Omega \setminus B).\end{aligned}$$

An ordered triple (A, B, C) of events in a probability space (Ω, P) is a **conjunctive fork** if

$$\begin{aligned}P(A \cap C|B) &= P(A|B) \cdot P(C|B), \\P(A \cap C|\Omega \setminus B) &= P(A|\Omega \setminus B) \cdot P(C|\Omega \setminus B), \\P(A|B) &> P(A|\Omega \setminus B), \\P(C|B) &> P(C|\Omega \setminus B).\end{aligned}$$

(implicit assumption $0 < P(B) < 1$)

An ordered triple (A, B, C) of events in a probability space (Ω, P) is a **conjunctive fork** if

$$\begin{aligned}P(A \cap C|B) &= P(A|B) \cdot P(C|B), \\P(A \cap C|\Omega \setminus B) &= P(A|\Omega \setminus B) \cdot P(C|\Omega \setminus B), \\P(A|B) &> P(A|\Omega \setminus B), \\P(C|B) &> P(C|\Omega \setminus B).\end{aligned}$$

(implicit assumption $0 < P(B) < 1$)

In contemporary language, $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C | \mathbf{1}_B$ and

$$\text{Cov}(\mathbf{1}_A, \mathbf{1}_B) > 0 \text{ and } \text{Cov}(\mathbf{1}_B, \mathbf{1}_C) > 0$$

An ordered triple (A, B, C) of events in a probability space (Ω, P) is a **conjunctive fork** if

$$\begin{aligned}P(A \cap C|B) &= P(A|B) \cdot P(C|B), \\P(A \cap C|\Omega \setminus B) &= P(A|\Omega \setminus B) \cdot P(C|\Omega \setminus B), \\P(A|B) &> P(A|\Omega \setminus B), \\P(C|B) &> P(C|\Omega \setminus B).\end{aligned}$$

(implicit assumption $0 < P(B) < 1$)

In contemporary language, $\mathbf{1}_A \perp\!\!\!\perp \mathbf{1}_C | \mathbf{1}_B$ and

$$\text{Cov}(\mathbf{1}_A, \mathbf{1}_B) > 0 \text{ and } \text{Cov}(\mathbf{1}_B, \mathbf{1}_C) > 0$$

where $\text{Cov}(\mathbf{1}_A, \mathbf{1}_B) = P(AB) - P(A)P(B)$

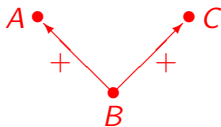
An ordered triple (A, B, C) of events in a probability space (Ω, P) is a **conjunctive fork** if

$$\begin{aligned}P(A \cap C|B) &= P(A|B) \cdot P(C|B), \\P(A \cap C|\Omega \setminus B) &= P(A|\Omega \setminus B) \cdot P(C|\Omega \setminus B), \\P(A|B) &> P(A|\Omega \setminus B), \\P(C|B) &> P(C|\Omega \setminus B).\end{aligned}$$

(implicit assumption $0 < P(B) < 1$)

In contemporary language, $1_A \perp\!\!\!\perp 1_C | 1_B$ and

$$\text{Cov}(1_A, 1_B) > 0 \text{ and } \text{Cov}(1_B, 1_C) > 0$$



where $\text{Cov}(1_A, 1_B) = P(AB) - P(A)P(B)$

Notation: $(A, B, C)_P$... the triple of events is a conjunctive fork

Notation: $(A, B, C)_P$... the triple of events is a conjunctive fork

Construction: having N finite and events $A_i, i \in N$, let

Notation: $(A, B, C)_P$... the triple of events is a conjunctive fork

Construction: having N finite and events A_i , $i \in N$, let

$$\{(i, j, k) \in N^3 : (A_i, A_j, A_k)_P\}.$$

Notation: $(A, B, C)_P$... the triple of events is a conjunctive fork

Construction: having N finite and events A_i , $i \in N$, let

$$\{(i, j, k) \in N^3 : (A_i, A_j, A_k)_P\}.$$

PROBLEM: Given a ternary relation \mathcal{R} on a ground set N decide whether it is *fork representable*, thus

$$(i, j, k) \in \mathcal{R} \Leftrightarrow (A_i, A_j, A_k)_P.$$

for some events A_i , $i \in N$.

Notation: $(A, B, C)_P$... the triple of events is a conjunctive fork

Construction: having N finite and events A_i , $i \in N$, let
 $\{(i, j, k) \in N^3: (A_i, A_j, A_k)_P\}$.

PROBLEM: Given a ternary relation \mathcal{R} on a ground set N decide whether it is *fork representable*, thus

$$(i, j, k) \in \mathcal{R} \Leftrightarrow (A_i, A_j, A_k)_P.$$

for some events A_i , $i \in N$.

In algebraic language,
solve a system of quadratic equations and inequalities.

$$\langle A, B \rangle \triangleq \text{Cov}(\mathbf{1}_A, \mathbf{1}_B) = P(AB) - P(A)P(B)$$

$$\langle A, B \rangle \triangleq \text{Cov}(\mathbf{1}_A, \mathbf{1}_B) = P(AB) - P(A)P(B)$$

Let $\sigma = (\sigma_{ij} \in \{-1, 0, 1\} : ij \in \binom{M}{2})$ be a pattern of signs indexed by the subsets ij with two elements.

$$\langle A, B \rangle \triangleq \text{Cov}(\mathbf{1}_A, \mathbf{1}_B) = P(AB) - P(A)P(B)$$

Let $\sigma = (\sigma_{ij} \in \{-1, 0, 1\} : ij \in \binom{N}{2})$ be a pattern of signs indexed by the subsets ij with two elements.

A simpler problem is easily solvable:

given any pattern σ , there exist events A_i , $i \in N$, s.t.

$$\sigma_{ij} = \text{sgn}\langle A_i, A_j \rangle, \quad ij \in \binom{N}{2}.$$

$$(A, B, C)_P \text{ iff } (C, B, A)_P$$

$(A, B, C)_P$ iff $(C, B, A)_P$

Lemma 1: $\langle A, B \rangle^2 \leq \langle A, A \rangle \langle B, B \rangle$, tight iff $1_A, 1_B$ lin. dependent

$(A, B, C)_P$ iff $(C, B, A)_P$

Lemma 1: $\langle A, B \rangle^2 \leq \langle A, A \rangle \langle B, B \rangle$, tight iff $1_A, 1_B$ lin. dependent

Lemma 2: If $1_A \perp 1_C | 1_B$ then $\langle A, C \rangle \langle B, B \rangle = \langle A, B \rangle \langle B, C \rangle$.

$$(A, B, C)_P \quad \text{iff} \quad (C, B, A)_P$$

Lemma 1: $\langle A, B \rangle^2 \leq \langle A, A \rangle \langle B, B \rangle$, tight iff $1_A, 1_B$ lin. dependent

Lemma 2: If $1_A \perp 1_C | 1_B$ then $\langle A, C \rangle \langle B, B \rangle = \langle A, B \rangle \langle B, C \rangle$.

Corollary 1: $(A, B, C)_P$ implies that A, B, C are nontrivial
 (thus $(A, A, A)_P, (B, B, B)_P$ and $(C, C, C)_P$)
 and any two are positively correlated
 (thus $(A, B, B)_P, (B, C, C)_P, (C, A, A)_P$)

$$(A, B, C)_P \quad \text{iff} \quad (C, B, A)_P$$

Lemma 1: $\langle A, B \rangle^2 \leq \langle A, A \rangle \langle B, B \rangle$, tight iff $1_A, 1_B$ lin. dependent

Lemma 2: If $1_A \perp 1_C | 1_B$ then $\langle A, C \rangle \langle B, B \rangle = \langle A, B \rangle \langle B, C \rangle$.

Corollary 1: $(A, B, C)_P$ implies that A, B, C are nontrivial
 (thus $(A, A, A)_P, (B, B, B)_P$ and $(C, C, C)_P$)
 and any two are positively correlated
 (thus $(A, B, B)_P, (B, C, C)_P, (C, A, A)_P$)

Corollary 2: $(A, B, C)_P$ and $(A, C, B)_P$ implies $B = C$

$$(A, B, C)_P \text{ iff } (C, B, A)_P$$

Lemma 1: $\langle A, B \rangle^2 \leq \langle A, A \rangle \langle B, B \rangle$, tight iff $1_A, 1_B$ lin. dependent

Lemma 2: If $1_A \perp 1_C | 1_B$ then $\langle A, C \rangle \langle B, B \rangle = \langle A, B \rangle \langle B, C \rangle$.

Corollary 1: $(A, B, C)_P$ implies that A, B, C are nontrivial
(thus $(A, A, A)_P, (B, B, B)_P$ and $(C, C, C)_P$)
and any two are positively correlated
(thus $(A, B, B)_P, (B, C, C)_P, (C, A, A)_P$)

Corollary 2: $(A, B, C)_P$ and $(A, C, B)_P$ implies $B = C$

Proof: $\langle A, C \rangle \langle B, B \rangle = \langle A, B \rangle \langle B, C \rangle$ and

$\langle A, B \rangle \langle C, C \rangle = \langle A, C \rangle \langle B, C \rangle$ combine to

$\langle B, C \rangle^2 = \langle B, B \rangle \langle C, C \rangle$, then $B = C$ □

To be fork representable, $\mathcal{R} \subseteq N^3$ must satisfy the symmetry

$$(i, j, k) \in \mathcal{R} \Rightarrow (k, j, i) \in \mathcal{R}$$

To be fork representable, $\mathcal{R} \subseteq N^3$ must satisfy the symmetry

$$(i, j, k) \in \mathcal{R} \Rightarrow (k, j, i) \in \mathcal{R}$$

by Corollary 2

$$(i, j, k) \in \mathcal{R} \text{ and } (i, k, j) \in \mathcal{R} \Rightarrow j = k$$

To be fork representable, $\mathcal{R} \subseteq N^3$ must satisfy the symmetry

$$(i, j, k) \in \mathcal{R} \Rightarrow (k, j, i) \in \mathcal{R}$$

by Corollary 2

$$(i, j, k) \in \mathcal{R} \text{ and } (i, k, j) \in \mathcal{R} \Rightarrow j = k$$

symmetry and \Leftrightarrow ... **betweenness**

To be fork representable, $\mathcal{R} \subseteq N^3$ must satisfy the symmetry

$$(i, j, k) \in \mathcal{R} \Rightarrow (k, j, i) \in \mathcal{R}$$

by Corollary 2

$$(i, j, k) \in \mathcal{R} \text{ and } (i, k, j) \in \mathcal{R} \Rightarrow j = k$$

symmetry and \Leftrightarrow ... **betweenness**

by Corollary 1

$$(i, j, k) \in \mathcal{R} \Rightarrow (i, i, i) \in \mathcal{R}, (j, j, j) \in \mathcal{R}, (k, k, k) \in \mathcal{R},$$

$$(i, j, k) \in \mathcal{R} \Rightarrow (i, j, j) \in \mathcal{R}, (j, k, k) \in \mathcal{R} \text{ and } (k, i, i) \in \mathcal{R}$$

To be fork representable, $\mathcal{R} \subseteq N^3$ must satisfy the symmetry

$$(i, j, k) \in \mathcal{R} \Rightarrow (k, j, i) \in \mathcal{R}$$

by Corollary 2

$$(i, j, k) \in \mathcal{R} \text{ and } (i, k, j) \in \mathcal{R} \Rightarrow j = k$$

symmetry and \Leftrightarrow ... **betweenness**

by Corollary 1

$$(i, j, k) \in \mathcal{R} \Rightarrow (i, i, i) \in \mathcal{R}, (j, j, j) \in \mathcal{R}, (k, k, k) \in \mathcal{R},$$

$$(i, j, k) \in \mathcal{R} \Rightarrow (i, j, j) \in \mathcal{R}, (j, k, k) \in \mathcal{R} \text{ and } (k, i, i) \in \mathcal{R}$$

Collecting the four implications ... **weak betweenness**

Lemma 3: $(A, B, A)_P$ iff $A = B$ is nontrivial

Lemma 3: $(A, B, A)_P$ iff $A = B$ is nontrivial

In the main problem assume $\mathcal{R} \subseteq N^3$ satisfies

Lemma 3: $(A, B, A)_P$ iff $A = B$ is nontrivial

In the main problem assume $\mathcal{R} \subseteq N^3$ satisfies
 $(i, j, i) \in \mathcal{R}$ iff $i = j$ for $i \in N$

Lemma 3: $(A, B, A)_P$ iff $A = B$ is nontrivial

In the main problem assume $\mathcal{R} \subseteq N^3$ satisfies

$(i, j, i) \in \mathcal{R}$ iff $i = j$ for $i \in N$

which excludes the trivial events and cloning

... the assumption of nondegeneracy

Lemma 3: $(A, B, A)_P$ iff $A = B$ is nontrivial

In the main problem assume $\mathcal{R} \subseteq N^3$ satisfies

$$(i, j, i) \in \mathcal{R} \text{ iff } i = j \text{ for } i \in N$$

which excludes the trivial events and cloning

... the assumption of nondegeneracy

In a weak betweenness \mathcal{R} for any 3-set ijk at most one of (i, j, k) ,
 (j, k, i) , (k, i, j) belongs to \mathcal{R}

Lemma 3: $(A, B, A)_P$ iff $A = B$ is nontrivial

In the main problem assume $\mathcal{R} \subseteq N^3$ satisfies

$$(i, j, i) \in \mathcal{R} \text{ iff } i = j \text{ for } i \in N$$

which excludes the trivial events and cloning

... the assumption of nondegeneracy

In a weak betweenness \mathcal{R} for any 3-set ijk at most one of (i, j, k) ,
 (j, k, i) , (k, i, j) belongs to \mathcal{R}

\mathcal{R} is called *solvable* if and only if the system

$$x_{ik} = x_{ij} + x_{jk} \quad \text{for } (i, j, k) \in \mathcal{R} \text{ pairwise distinct,}$$

has a solution with all involved x_{ij} positive.

Theorem

Under nondegeneracy, $\mathcal{R} \subseteq N^3$ is fork representable iff it is a solvable weak betweenness.

Theorem

Under nondegeneracy, $\mathcal{R} \subseteq N^3$ is fork representable iff it is a solvable weak betweenness.

In general, \mathcal{R} must be a 'regular' weak betweenness. Then a *quotient* Q of \mathcal{R} is constructed. It is a weak betweenness that satisfies the nondegeneracy condition. \mathcal{R} is fork representable iff Q is solvable.

Theorem

Under nondegeneracy, $\mathcal{R} \subseteq N^3$ is fork representable iff it is a solvable weak betweenness.

In general, \mathcal{R} must be a 'regular' weak betweenness. Then a *quotient* Q of \mathcal{R} is constructed. It is a weak betweenness that satisfies the nondegeneracy condition. \mathcal{R} is fork representable iff Q is solvable.

The conditions can be verified in time polynomial in $|N|$.

NECESSITY

NECESSITY

Assume nondegeneracy and \mathcal{R} to be fork representable by $A_i, i \in N$.

NECESSITY

Assume nondegeneracy and \mathcal{R} to be fork representable by A_i , $i \in N$.

For i, k different, participating in a fork, let

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_i, A_k \rangle} .$$

NECESSITY

Assume nondegeneracy and \mathcal{R} to be fork representable by A_i , $i \in N$.

For i, k different, participating in a fork, let

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_i, A_k \rangle}.$$

By Lemma 1, $x_{ik} > 0$. By Lemma 2, \mathcal{R} is solvable since

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_j, A_j \rangle^{1/2}}{\langle A_i, A_j \rangle} - \frac{\langle A_j, A_j \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_j, A_k \rangle} = x_{ij} + x_{jk}$$

NECESSITY

Assume nondegeneracy and \mathcal{R} to be fork representable by A_i , $i \in N$.

For i, k different, participating in a fork, let

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_i, A_k \rangle}.$$

By Lemma 1, $x_{ik} > 0$. By Lemma 2, \mathcal{R} is solvable since

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_j, A_j \rangle^{1/2}}{\langle A_i, A_j \rangle} - \frac{\langle A_j, A_j \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_j, A_k \rangle} = x_{ij} + x_{jk}$$

for $(A_i, A_j, A_k)_P$.

NECESSITY

Assume nondegeneracy and \mathcal{R} to be fork representable by A_i , $i \in N$.

For i, k different, participating in a fork, let

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_i, A_k \rangle}.$$

By Lemma 1, $x_{ik} > 0$. By Lemma 2, \mathcal{R} is solvable since

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_j, A_j \rangle^{1/2}}{\langle A_i, A_j \rangle} - \frac{\langle A_j, A_j \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_j, A_k \rangle} = x_{ij} + x_{jk}$$

for $(A_i, A_j, A_k)_P$.

SUFFICIENCY

NECESSITY

Assume nondegeneracy and \mathcal{R} to be fork representable by A_i , $i \in N$.

For i, k different, participating in a fork, let

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_i, A_k \rangle}.$$

By Lemma 1, $x_{ik} > 0$. By Lemma 2, \mathcal{R} is solvable since

$$x_{ik} = \ln \frac{\langle A_i, A_i \rangle^{1/2} \langle A_j, A_j \rangle^{1/2}}{\langle A_i, A_j \rangle} - \frac{\langle A_j, A_j \rangle^{1/2} \langle A_k, A_k \rangle^{1/2}}{\langle A_j, A_k \rangle} = x_{ij} + x_{jk}$$

for $(A_i, A_j, A_k)_P$.

SUFFICIENCY

P is constructed on \mathbb{Z}_2^N explicitly, can be arbitrarily close to the uniform distribution; Fourier-Stieltjes transform of P is related to solvability + few other tricks

Given a pattern of signs $(\sigma_{ij} : ij \in \binom{N}{2})$ and a family $\{(i, j, k)\}$,
represent them simultaneously by events.

Given a pattern of signs $(\sigma_{ij} : ij \in \binom{N}{2})$ and a family $\{(i, j, k)\}$,
represent them simultaneously by events.

For $P > 0$ likely in reach (including primary decomposition),
for $P \geq 0$ open ??

Given a pattern of signs $(\sigma_{ij} : ij \in \binom{N}{2})$ and a family $\{(i, j, k)\}$,
represent them simultaneously by events.

For $P > 0$ likely in reach (including primary decomposition),
for $P \geq 0$ open ??

Given a family $\{(i, j, k)\}$, represent it by arbitrary variables
($P > 0$ or $P \geq 0$) ??

Given a pattern of signs $(\sigma_{ij} : ij \in \binom{N}{2})$ and a family $\{(i, j, k)\}$,
 represent them simultaneously by events.

For $P > 0$ likely in reach (including primary decomposition),
 for $P \geq 0$ open ??

Given a family $\{(i, j, k)\}$, represent it by arbitrary variables
 $(P > 0 \text{ or } P \geq 0)$??

Gaussian case is likely not difficult.