The Logic of Dependence and Independence

Jouko Väänänen\textsuperscript{1}

Helsinki and Amsterdam

Nov 2011

\textsuperscript{1}Supported by the ESF Eurocores LogICCC program project \textit{Logic for Interaction} (LINT).
Dependence and Independence

- Random variables
- Equations, linear dependence, algebraic dependence
- Causality
- Game theory
- Genetics
- Data mining

Are there logical properties of dependence and independence concepts that underlie all these different disciplines?
Dependence and independence (as they occur in statistics, experimental science, computer science, etc)

- Can be treated as **atoms** in logic, like their cousin **identity**.
- Are **logical** concepts.
- Can be **axiomatized**.
- Give rise to an emerging new fundamental (logical) theory.
The new atomic formula

\[(\vec{x}, \vec{y})\]

has the intuitive interpretation “the values of the variables \(\vec{y}\) are completely determined by the values of the variables \(\vec{x}\)”.
The new atomic formula

\[=(\vec{x}, \vec{y})\]

has the intuitive interpretation “the values of the variables \(\vec{y}\) are completely determined by the values of the variables \(\vec{x}\)”.  

Armstrong’s Axioms:

1. \[=(\vec{x}, \vec{x}).\]

2. If \[=(\vec{y}, \vec{x})\] and \(\vec{y} \subseteq \vec{z}\), then \[=(\vec{z}, \vec{x}).\]

3. If \(\vec{y}\) is a permutation of \(\vec{z}\), \(\vec{u}\) is a permutation of \(\vec{x}\), and \[=(\vec{z}, \vec{x}),\] then \[=(\vec{y}, \vec{u}).\]

4. If \[=(\vec{y}, \vec{z})\] and \[=(\vec{z}, \vec{x}),\] then \[=(\vec{y}, \vec{x}).\]
We use

$$\vec{x} \perp \vec{y}$$

to denote an atomic formula with the intuitive interpretation “the values of the variables $\vec{x}$ are completely independent of the values of the variables $\vec{y}$”.
We use

\[ \bar{x} \perp \bar{y} \]

to denote an atomic formula with the intuitive interpretation “the values of the variables \( \bar{x} \) are completely independent of the values of the variables \( \bar{y} \)”.  

**Axioms (Geiger-Paz-Pearl):**

1. If \( \bar{y} \perp \bar{x} \), then \( \bar{x} \perp \bar{y} \).
2. If \( \bar{y} \perp \bar{x} \) and \( \bar{z} \subseteq \bar{y} \), then \( \bar{z} \perp \bar{x} \).
3. If \( \bar{y} \) is a permutation of \( \bar{z} \), \( \bar{u} \) is a permutation of \( \bar{x} \), and \( \bar{z} \perp \bar{x} \), then \( \bar{y} \perp \bar{u} \).
4. If \( \bar{y} \perp \bar{z} \) and \( \bar{y} \bar{z} \perp \bar{x} \), then \( \bar{y} \perp \bar{z} \bar{x} \).

Note: \( = (\bar{x}) \) is equivalent to \( \bar{x} \perp \bar{x} \).
We use

$$\vec{x} \subseteq \vec{y}$$

to denote an atomic formula with the intuitive interpretation “every value of $\vec{x}$ occurs as a value of $\vec{y}$”.
We use

\[ \vec{x} \subseteq \vec{y} \]

to denote an atomic formula with the intuitive interpretation "every value of \( \vec{x} \) occurs as a value of \( \vec{y} \)."

- Axiomatized by Casanova-Fagin-Papadimitriou.
- Mitchell, Chandra-Vardi: Inclusion and dependence atoms together can be axiomatized.
Exclusion atom

We use

$$\vec{x} \mid \vec{y}$$

to denote an atomic formula with the intuitive interpretation “no value of $\vec{x}$ occurs as a value of $\vec{y}$”. 
Exclusion atom

We use

\[ \vec{x} \mid \vec{y} \]

to denote an atomic formula with the intuitive interpretation “no value of \( \vec{x} \) occurs as a value of \( \vec{y} \)”.

- Axiomatized by Casanova-Vidal 1983.
Conditional independence

We use

\[ \vec{x} \perp_{\vec{z}} \vec{y} \]

to denote the atomic formula with the intuitive interpretation “the values of \( \vec{x} \) are independent of the values of \( \vec{y} \), if the value of \( \vec{z} \) is kept fixed”.
Conditional independence

We use

\[ \vec{x} \perp_{\vec{z}} \vec{y} \]

to denote the atomic formula with the intuitive interpretation “the values of \( \vec{x} \) are independent of the values of \( \vec{y} \), if the value of \( \vec{z} \) is kept fixed”.

- Can be axiomatized.
- \( \equiv(\vec{x}, \vec{y}) \) is equivalent to \( \vec{y} \perp_{\vec{x}} \vec{y} \).
Conclusion:

- Contemplation on dependence and independence notions gives rise to a variety of purely logical atoms, as well as axioms that completely describe them.
Whatever dependence/independence atoms we have, we can coherently add logical operations \(\land\), \(\lor\), \(\forall\) and \(\exists\).

In front of the atoms can also use \(\neg\).

Conservative extension of classical logic.

Subtlety: the logical operations have variants. The differences do not manifest themselves in first order logic, only in connection with the new atoms.
Semantics

- A **team** is a set of assignments (or a table of data, database, etc).
- The point (W. Hodges): The dependence - independence phenomena do not manifest themselves in the presence of only one assignment.
- With teams we can give meaning to formulas involving \( \land, \lor, \forall, \exists, \neg \) and the new atoms.
Definition

A team $X$ satisfies the atom $≡(\vec{x}, \vec{y})$ if

$$\forall s, s' \in X (s(\vec{x}) = s'(\vec{x}) \rightarrow s(\vec{y}) = s'(\vec{y})).$$
**Definition**

A team $X$ satisfies the atom $=(\vec{x}, \vec{y})$ if

$$\forall s, s' \in X (s(\vec{x}) = s'(\vec{x}) \rightarrow s(\vec{y}) = s'(\vec{y})).$$

**Example**

$X =$ scientific data about dropping lead balls in Pisa. $X$ satisfies $=(\text{height}, \text{time})$

if in any two drops from the same height the times are the same.
Definition

A team $X$ satisfies the atomic formula $\vec{y} \perp_{\vec{x}} \vec{z}$ if for all $s, s' \in X$ such that $s(\vec{x}) = s'(\vec{x})$ there exists $s'' \in X$ such that $s''(\vec{x}) = s(\vec{x})$, $s''(\vec{y}) = s(\vec{y})$, and $s''(\vec{z}) = s'(\vec{z})$.

Similarly, inclusion $\vec{x} \subseteq \vec{y}$ and exclusion $\vec{x} \mid \vec{y}$.
Definition

A team $X$ satisfies the atomic formula $\vec{y} \perp_{\vec{x}} \vec{z}$ if for all $s, s' \in X$ such that $s(\vec{x}) = s'(\vec{x})$ there exists $s'' \in X$ such that $s''(\vec{x}) = s(\vec{x})$, $s''(\vec{y}) = s(\vec{y})$, and $s''(\vec{z}) = s'(\vec{z})$.

Similarly inclusion $\vec{x} \subseteq \vec{y}$ and exclusion $\vec{x} \mid \vec{y}$.

Example

$X$ = scientific experiment setup concerning dropping lead balls of different sizes in Pisa. $X$ should satisfy

weight $\perp_{\text{size}}$ height

i.e. for any two drops of a ball of a fixed size, also a drop with that size, weight from the first and height from the second, is performed.
Definition

A team $X$ satisfies $\phi \lor \psi$ if $X = Y \cup Z$ such that $Y$ satisfies $\phi$ and $Z$ satisfies $\psi$. 
Dependence and independence concepts
The fundamental logics
Position on the map of logics
Complete axiomatization
Applications

Pietro Galliani (Amsterdam), others
Pietro Galliani (Amsterdam), others
Infinity

- $\forall x \forall y \exists z ((z, y) \land \lnot z = x)$ characterizes infinity.
- Cannot axiomatize logical consequence. (Even worse: logical consequence encodes the CH.)
- Add intuitionistic implication, get full second order logic.
- **Can** axiomatize first order consequences.
The rules

- Natural deduction of classical logic, **but Disjunction Elimination Rule** and **Negation Introduction Rule** only for first order formulas.
- Weak Disjunction Rule: From $\psi \vdash \theta$ conclude $\phi \lor \psi \vdash \phi \lor \theta$.
- Dependence Introduction Rule:
  $$\exists y \forall x \phi(x, y, \vec{z}) \vdash \forall x \exists y (=(\vec{z}, y) \land \phi(x, y, \vec{z})).$$
- Dependence Distribution rule
- Dependence Elimination Rule
- **Completeness Theorem** (Model theoretic proof).
Dependence distribution: let

\[ A = \exists y_1 \ldots \exists y_n \bigwedge_{1 \leq j \leq n} \left( (\bar{z}^j, y_j) \land C \right), \]

\[ B = \exists y_{n+1} \ldots \exists y_{n+m} \bigwedge_{n+1 \leq j \leq n+m} \left( (\bar{z}^j, y_j) \land D \right). \]

where \( C \) and \( D \) are quantifier-free formulas without dependence atoms, and \( y_i \), for \( 1 \leq i \leq n \), does not appear in \( B \) and \( y_i \), for \( n + 1 \leq i \leq n + m \), does not appear in \( A \). Then,

\[
\frac{A \lor B}{\exists y_1 \ldots \exists y_{n+m} \bigwedge_{1 \leq j \leq n+m} \left( (\bar{z}^j, y_j) \land (C \lor D) \right)}
\]

Note that the logical form of this rule is:

\[
\frac{\exists \bar{y} \bigwedge_{1 \leq j \leq n} \left( (\bar{z}^j, y_j) \land C \right) \lor \exists \bar{y}' \bigwedge_{n+1 \leq j \leq n+m} \left( (\bar{z}^j, y_j) \land D \right)}{\exists \bar{y} \exists \bar{y}' \bigwedge_{1 \leq j \leq n+m} \left( (\bar{z}^j, y_j) \land (C \lor D) \right)}
\]
**Dependence elimination:**

\[
\forall \bar{x}_0 \exists \bar{y}_0 (\bigwedge_{1 \leq j \leq k} = (\bar{w}^{i_j}, y_{0,i_j}) \land B(\bar{x}_0, \bar{y}_0)),
\]

\[
\forall \bar{x}_0 \exists \bar{y}_0 (B(\bar{x}_0, \bar{y}_0) \land \forall \bar{x}_1 \exists \bar{y}_1 (B(\bar{x}_1, \bar{y}_1) \land \bigwedge_{(\bar{w}_0^p, y_{0,p}) \in S} (\bar{w}_0^p = \bar{w}_1^p \rightarrow y_{0,p} = y_{1,p}))),
\]

where \( \bar{x}_i = (x_{l_1}, \ldots, x_{l_m}) \) and \( \bar{y}_l = (y_{l_1}, \ldots, y_{l_n}) \) for \( l \in \{0, 1\} \) (\( \bar{w}_0^p \) and \( \bar{w}_1^p \) are related analogously), and the variables in \( \bar{w}^{i_j} \) are contained in the set

\[\{x_{0,1}, \ldots, x_{0,m}, y_{0,1}, \ldots, y_{0,i_j-1}\}.
\]

Furthermore, the set \( S \) contains the conjuncts of

\[\bigwedge_{1 \leq j \leq k} = (\bar{w}^{i_j}, y_{0,i_j}),
\]

and the dependence atom \( = (x_{0,1}, \ldots, x_{0,m}, y_{0,p}) \) for each of the variables \( y_{0,p} \) (1 \( \leq p \leq n \)) such that \( y_{0,p} \notin \{y_{0,i_1}, \ldots, y_{0,i_k}\} \).
Dependence Distribution Rule

1. Given $\epsilon, x, y$ and $f$.

2. If $\epsilon > 0$, then there is $\delta > 0$ depending only on $\epsilon$ such that if $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.

3. Therefore, there is $\delta > 0$ depending only on $\epsilon$ such that if $\epsilon > 0$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$. 
1 Assume that for every $x$ and every $\epsilon > 0$ there is $\delta > 0$ depending only on $\epsilon$ such that for all $y$, if $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.

2 Therefore, for every $x$ and every $\epsilon > 0$ there is $\delta > 0$ such that for all $y$, if $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$, and moreover, for another $x'$ and $\epsilon' > 0$ there is $\delta' > 0$ such that for all $y'$, if $|x' - y'| < \delta'$, then $|f(x') - f(y')| < \epsilon$ and if $\epsilon = \epsilon'$, then $\delta = \delta'$.
Social choice

- Individual voters are the **variables**.
- The **values** of these variables are the preference relations of the individuals $x_i$.
- An **assignment** = a profile.
- The social choice function is just another variable $y$.
- Independence from irrelevant alternatives:
  $$=(P_{a,b}(x_1), \ldots, P_{a,b}(x_n), P_{a,b}(y)).$$
- $x_i$ is a dictator: $=(x_i, y)$.
- Arrow’s axioms invoke dependence only, but the **proof** of Arrow’s theorem depends on assumptions that invoke independence assumptions about the behavior of the electorate.
The proof of Arrow’s theorem assumes—seemingly—that the social welfare function is defined for all profiles.

**Enough**: the domain of the social welfare function manifests independence of the voters from each other in the sense that profiles where voters change their preferences (as needed in the proof) are also possible.

**Arrow’s Paradox**: Unlimited freedom leads to the situation that dictatorship is the only way to satisfy Arrow's axioms.
Samson Abramsky: “Relational Hidden Variables and Non-Locality”.

Team = a set of observations. Experiments \( q_1, \ldots, q_n \). Each has an input and an output. Input of experiment \( q_i \) denoted \( x_i \), and the output \( y_i \). After \( m \) rounds of making the experiments \( q_1, \ldots, q_m \) we have the data i.e. the team

\[
X = \begin{array}{cccc}
\begin{array}{c}
X_1 \\
Y_1 \\
\vdots \\
X_n \\
Y_n \\
\end{array} & \begin{array}{c}
\begin{array}{c}
a_1^1 \\
a_1^2 \\
\vdots \\
a_1^n \\
\end{array} \\
\begin{array}{c}
b_1^1 \\
b_1^2 \\
\vdots \\
b_1^n \\
\end{array} \\
\end{array} & \ldots & \begin{array}{c}
\begin{array}{c}
a_2^1 \\
a_2^2 \\
\vdots \\
a_2^n \\
\end{array} \\
\begin{array}{c}
b_2^1 \\
b_2^2 \\
\vdots \\
b_2^n \\
\end{array} \\
\end{array} & \ldots & \begin{array}{c}
\begin{array}{c}
a_n^1 \\
a_n^2 \\
\vdots \\
a_n^n \\
\end{array} \\
\begin{array}{c}
b_n^1 \\
b_n^2 \\
\vdots \\
b_n^n \\
\end{array} \\
\end{array}
\end{array}
\]
Team X supports **strong determinism** if it satisfies

\[ (x_i, y_i) \]

for all \( i = 1, \ldots, n \).

Team X supports **weak determinism** if it satisfies

\[ (x_1, \ldots, x_n, y_i) \]

for all \( i = 1, \ldots, n \).
Empirical models

A dependence (and independence) logic formula $\phi(\vec{x}, \vec{y})$ describes a “relational model”.

**Example**

$$\exists X \exists Y \exists a \exists b (x_1 = X \land x_2 = Y \land (y_1 = a \lor y_1 = b) \land (y_2 = a \lor y_2 = b))$$

says that in the model there are two experiments, both have a fixed input ($X$ and $Y$), with two possible outcomes ($a$ or $b$) for either experiment. An example of such a team would be

$$X = \begin{array}{cccc}
  x_1 & x_2 & y_1 & y_2 \\
  X & a & Y & b \\
  X & b & Y & a \\
\end{array}$$
### Hidden variables

A **hidden variable model** is of the form

\[
Y = \begin{array}{cccccc}
  x_1 & y_1 & \ldots & x_n & y_n & z \\
  a_1 & b_1 & \ldots & a_n & b_n & \gamma^1 \\
  a_2 & b_2 & \ldots & a_n & b_n & \gamma^2 \\
  \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\
  a_m & b_m & \ldots & a_n & b_n & \gamma^m \\
\end{array}
\]

where the \( \gamma^i \) are values of a hidden variable \( z \).

A hidden variable model \( Y \) **satisfies** the relational model \( X \) (i.e. \( X \) is **realized** by \( Y \)) if

\[
X = \exists z Y,
\]

that is

\[
s \in X \iff \exists s' \in Y (s'(x_1) = s(x_1) \land s'(y_1) = s(y_1) \land \ldots \\
s'(x_n) = s(x_n) \land s'(y_n) = s(y_n)).
\]
A team $X$ is said to support **single-valuedness of the hidden variable** $z$ if $z$ has only one value in the team. We can express this with the formula

$$=(z).$$
Outcome-independence

An empirical team $X$ is said to support **outcome-independence** if the following holds: Suppose the team $X$ has two measurement-outcome combinations $s$ and $s'$ with the same total input data $\vec{x}$ and the same hidden variable $z$, i.e. $s(\vec{x}) = s'(\vec{x})$ and $s(z) = s'(z)$. We demand that output $s(y_i)$ should occur as an output also if the outputs $s(\{y_j : j \neq i\})$ are changed to $s'(\{y_j : j \neq i\})$.

We can express outcome-independence with the formula

$$y_i \perp_{\vec{x},z} \{y_j : j \neq i\}.$$
Other

- No-signaling.
- Independence of the hidden variable
- Parameter independence
- Locality
No-Go Results

- **Einstein-Podolsky-Rosen result**: There is an empirical model (team) which cannot be realized by any hidden variable models satisfying single-valuedness and outcome-independence.

- **Other**: Greenberger-Horne-Zeilinger result, Hardy paradox, Kochen-Specker Theorem.
The emergent logic of dependence and independence concepts provides a common mathematical and conceptual basis not only for statistics, algebra, causality, game theory, and data mining, but also for phenomena such as Arrow’s Paradox and the no-go-results of Quantum Mechanics.

We can find fundamental principles governing this logic.

Important parts can be completely axiomatized, other parts are manifestly beyond the reach of axiomatization.