Dependence logic

Jouko Väänänen

University of Helsinki and University of Amsterdam

Auckland, April 2016

Everyday language

- I will park the car next to the lamp post depending only on the day of the week.
- I will park the car next to the lamp post independently of the day of the week.
- The time of descent of the ball depends only on the height of the drop.
- The time of descent of the ball is independent of the weight of the ball.

Database dependencies

functional dependency:	`name`	$\rightarrow$	`phone`
inclusion:	`teacher`	$\subseteq$	`personnel`
independence:	`salary`	$\perp$	`gender`
conditional independence:	`salary`	$\perp_{\texttt{job}}$	`gender`

Statistical dependencies

smoking causes lung cancer

diabetes is independent of gender

mutation-1 is independent of mutation-2, if we fix gender

Algebraic dependencies

$2x + 5y$, $x + y$ and $3x + 6y$ are linearly dependent$/\mathbb{Q}$

π and e are algebraically independent$/\mathbb{Q}$

Biological dependencies

The biological sex depends on the XY-chromosomes

The biological sex is independent of nutrition

An allele of a gene is inherited independently from mother and father

Economics

- Arrow's Theorem: If a set of voting profiles satisfies independence ("freedom of expression") assumptions, and the axioms of Pareto efficiency and dependence on only relevant alternatives, then in view of this data there is a dictator.

Game theory

- If a player is following a fixed strategy, his or her moves completely depend on the opponent's moves.
- Imperfect information game: A player is required to make some moves independently of what others played.

Physics

- Quantum physics provides a rich field of highly non-trivial dependence and independence concepts. Some of the most fundamental questions of quantum physics are about independence of outcomes of experiments.

Dependence logic

- Dependence and independence concepts are ubiquitous.
- We show that there is a common core.
- We develop a logic for the study of the common core, and call it the dependence logic.

Dependence and independence as atoms

- Dependence atom $=(x, y)$, "y depends only on x". (Database theory notation: $x \to y$.)
- Independence atom $x \perp y$, "x and y are independent of each other".
- Relativized independence atom $x \perp_z y$, "x and y are independent of each other, if z is kept fixed".
- Inclusion atom $x \subseteq y$, "values of x occur also as values of y".

Armstrong's Axioms

1.	Identity rule:	$=\!(x, x)$.
2.	Symmetry Rule:	If $=\!(xt, yr)$, then $=\!(tx, ry)$.
3.	Weakening Rule:	If $=\!(x, yr)$, then $=\!(xt, y)$.
4.	Transitivity Rule:	If $=\!(x, y)$ and $=\!(y, r)$, then $=\!(x, r)$.

Geiger-Paz-Pearl axioms

1.	Empty set rule:	$x \perp \emptyset$.
2.	Symmetry Rule:	If $x \perp y$, then $y \perp x$.
3.	Weakening Rule:	If $x \perp yr$, then $x \perp y$.
4.	Exchange Rule:	If $x \perp y$ and $xy \perp r$, then $x \perp yr$.

Axioms of relative independence

Definition

The **axioms** of the relative independence atom are:

1. $y \perp_x y$ entails $y \perp_x z$ (Constancy Rule)
2. $x \perp_x y$ (Reflexivity Rule)
3. $z \perp_x y$ entails $y \perp_x z$ (Symmetry Rule)
4. $yy' \perp_x zz'$ entails $y \perp_x z$. (Weakening Rule)
5. If z' is a permutation of z, x' is a permutation of x, y' is a permutation of y, then $y \perp_x z$ entails $y' \perp_{x'} z'$. (Permutation Rule)
6. $z \perp_x y$ entails $yx \perp_x zx$ (Fixed Parameter Rule)
7. $x \perp_z y \wedge u \perp_{zx} y$ entails $u \perp_z y$. (First Transitivity Rule)
8. $y \perp_z y \wedge zx \perp_y u$ entails $x \perp_z u$ (Second Transitivity Rule)

Semantics of dependence and independence

Each application area has its own semantics:

- Databases.
- Tables of values of random variables.
- Bayesian networks.
- Vector spaces, algebraically closed fields.
- Sets of (voting) profiles.
- Sets of plays in a game (*the original motivation*)
- Tables of observations in experimental science.
- Probability tables in quantum physics.
- State of information (under uncertainty).
- Randomness, encryption, etc

Team semantics—a general multi-purpose concept

Definition

A *team* is a pair (X, τ), where X is a set and τ is a function such that

1. $dom(\tau) = X$,
2. If $i \in X$, then $\tau(i)$ is a function which maps elements of a fixed set of attributes (variables) to a fixed domain.

Loosely speaking, a team is just a table of values of a fixed collection of attributes (a.k.a. variables).

Typical team: X is a set of individuals, $\tau(i)$ is an assignment of alleles to genes and phenotypes to traits.

$$X = \left\{ \begin{array}{|c|c|} \hline \text{Genes} & \text{Traits} \\ \hline \cdots & \cdots \\ \hline \vdots & \vdots \\ \hline \cdots & \cdots \\ \hline \end{array} \right.$$

(∼20k genes)	BRCA	**RECQL**	Pol/Can	Cancer
...AACTTCGAGGCTTACCGCTG...	1	0	P	0
...AAGGTCGATGCTCACCGATG...	1	1	P	1
⋮ (25k cases)	⋮	⋮	⋮	⋮
...AACGTCTATGCTCACCGATG...	1	1	C	1

Semantics of the identity atom

Definition

A team X satisfies the atom $x = y$ if

Any two rows have the same value for x and y.

Example

X = scientific data about dropping iron balls in Pisa with two timing methods `time1` and `time2`. X satisfies

$$\texttt{time1} = \texttt{time2}$$

if the two times `time1` and `time2` agree in all drops.

Semantics of the dependence atom

Definition

A team X satisfies the atom $=(x, y)$ (i.e. $x \to y$) if

Any two rows with the same value for x has also the same value for y.

Example

$X =$ scientific data about dropping iron balls in Pisa. X satisfies

$$=(\text{height}, \text{time})$$

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

Semantics of the independence atom

Definition

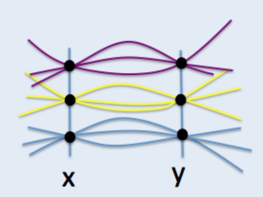
A team X satisfies the atomic formula $y \perp z$ if for all $s, s' \in X$ there exists $s'' \in X$ such that $s''(y) = s(y)$, and $s''(z) = s'(z)$.

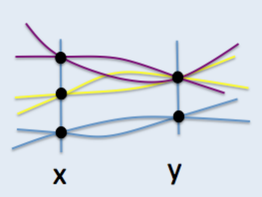
Example

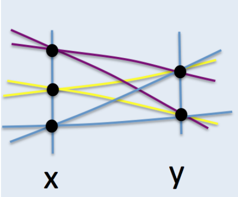
X = scientific experiment concerning dropping iron balls of a fixed size from a fixed height in Pisa. X satisfies

$$\text{weight} \perp \text{time}$$

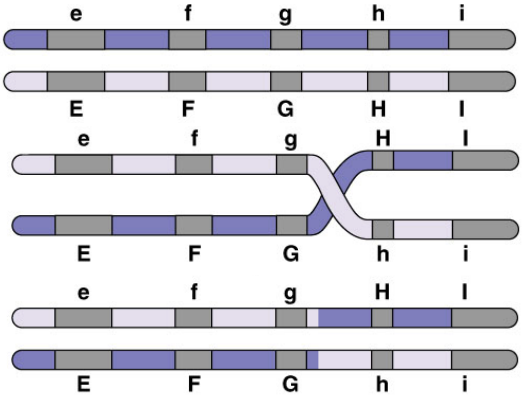
if for any two drops of a ball also a drop, with weight from the first and time from the second, is observed.







$$x = y \qquad =\!(x, y) \qquad x \perp y$$


e
f
g
h
i
E
F
G
H
I
e
f
g
H
I
E
F
G
h
i
e
f
g
H
I
E
F
G
h
i

A Completeness Theorem

Theorem (Armstrong)

If T is a finite set of dependence atoms of the form $=(u, v)$ for various u and v, then TFAE:

1. $=(x, y)$ *follows from T according to the above rules.*
2. *Every team that satisfies T also satisfies $=(x, y)$.*

Note: Holds also for other semantics!

A Completeness Theorem

Theorem (Geiger-Paz-Pearl)

If T is a finite set of independence atoms of the form $t \perp r$ for various t and r, then TFAE:

1. $x \perp y$ *follows from T according to the above rules*
2. *Every team that satisfies T also satisfies* $x \perp y$.

Note: Holds also for other semantics!

A logic of dependence?

All men are mortal.	$\forall x(A(x) \to B(x))$
Socrates is a man.	$A(c)$
Socrates is mortal.	$B(c)$

(∼20k genes)	BRCA	**RECQL**	Pol/Can	Cancer
...AACTTCGAGGCTTACCGCTG...	1	0	P	0
...AAGGTCGATGCTCACCGATG...	1	1	P	1
⋮ (25k cases)	⋮	⋮	⋮	⋮
...AACGTCTATGCTCACCGATG...	1	1	C	1

(Germline RECQL mutations are associated with cancer susceptibility, Cybulski et al (2015))

Logical operations

- Whatever dependence/independence atoms we have, we can coherently add logical operations "and", "or", "for all" and "exists".
- In front of the first order atoms can also use negation.
- Conservative extension of classical logic.
- Discovery of more complicated patterns than mere dependency constraints.

Disjunction

Definition

A team X satisfies $\phi \vee \psi$ if $X = Y \cup Z$ such that Y satisfies ϕ and Z satisfies ψ.

Example

Example

A directed graph contains a cycle (or an infinite path) iff it satisfies
$\exists x \exists y(y \subseteq x \wedge xEy)$

Logic of dependence and independence

Definition

1. Dependence logic is the extension of first order logic obtained by adding the dependence atoms $=(x, y)$. (V. 2007)
2. Independence logic is the extension of first order logic obtained by adding the independence atoms $x \perp y$. (Grädel-V. 2010)
3. Inclusion logic is the extension of first order logic obtained by adding the inclusion atoms $x \subseteq y$. (Galliani 2012)

Logic of dependence and independence

Lemma

- *Dependence logic is downward closed: If a team has a property expressed by a dependence logic formula, every subteam has the property, too.*
- *Inclusion logic is closed under unions: if two team have a property expressed by an inclusion logic formula, their union has the property, too.*

The fundamental characterizations

Theorem

- *Dependence logic = downward closed non-deterministic polynomial time. (Kontinen-V. 2009)*
- *Independence logic = non-deterministic polynomial time. (Galliani 2012)*
- *Inclusion logic = polynomial time on ordered models. (Galliani-Hella 2013)*

The rules

Definition

- 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 $\phi \vee \psi \vdash \phi \vee \theta$.
- Dependence Introduction Rule:
  $\exists y \forall x \phi(x, y, \vec{z}) \vdash \forall x \exists y (=\!(\vec{z}, y) \wedge \phi(x, y, \vec{z}))$.
- Dependence Distribution rule
- Dependence Elimination Rule

Theorem (Completeness Theorem)

- *The above axioms and rules are complete for first order consequences from dependence logic assumptions. (Kontinen-V. 2011)*
- *The same for independence logic. (Hannula 2015)*

Social choice

Suppose we have n voters $x_1, \ldots, x_n$, each giving his or her (linear) preference quasi-order $<_{x_i}$ on some finite set A of alternatives. We call such sequences $p_1, \ldots, p_n$ **profiles**.

Let us denote the social choice by y, which is likewise a preference order $<_y$.

Naturally we assume

$$=\!(x_1, \ldots, x_n, y).$$

1. A team is **Paretian** if the team satisfies the first order formula:

$$(a <_{x_1} b \wedge \ldots \wedge a <_{x_n} b) \rightarrow a <_y b,$$

   for all $a, b \in A$. Note that this means that every individual row satisfies the formula.

2. A team is **dictatorial** if in the team

$$x_1 = y \vee_B \ldots \vee_B x_n = y.$$

3. A team **respects independence of irrelevant alternatives** if it satisfies for all $a, b \in A$:

$$=\!(\{a <_{x_1} b, \ldots, a <_{x_n} b\}, \{a <_y b\}).$$

   Note that this is a Boolean dependence atom.

4. A team supports **voting independence**, if it satisfies for all i:

$$x_i \perp \{x_j : j \neq i\}.$$

Definition

We introduce a new *universality* *atom* $\forall(x_1, \ldots, x_n)$ with the intuitive meaning that any combination of values (in the given domain) for $x_1, \ldots, x_n$ is possible. A team X satisfies

$$\forall(x_1, \ldots, x_n),$$

if for every $a_1, \ldots, a_n \in M$ there is $s \in X$ such that $s(x_1) = a_1, \ldots, s(x_n) = a_n$.

Freedom

$$\forall(x_1) \wedge \ldots \wedge \forall(x_n)$$

The **freedom of choice** assumption. Together with voting independence it implies that all patterns of voting can arise.

Theorem (Arrow 1963)

Voting independence, freedom of choice, Pareto and respect of independence of irrelevant alternatives together imply dictatorship. In symbols,

$$
\begin{array}{l}
\{=(x_1,\ldots,x_n,y), \\
\bigwedge_{a,b\in A}((a\leq_{x_1} b\wedge\ldots\wedge a\leq_{x_n} b)\to a\leq_y b), \\
\bigwedge_{a,b\in A} =(\{a\leq_{x_1} b,\ldots,a\leq_{x_n} b\},\{a\leq_y b\}), \\
\forall(x_1),\ldots,\forall(x_n),\bigwedge_{i=1}^n x_i\perp\{x_j : j\neq i\}\} \\
\qquad\qquad\qquad \models x_1=y\vee_B\ldots\vee_B x_n=y.
\end{array}
$$

Observational teams

One of the intuitions behind the concept of a team is a set of observations, such as readings of physical measurements. Let us consider a **system** consisting of experiments

$$q_1, ..., q_n.$$

Each experiment has an input x_i and an output y_i.

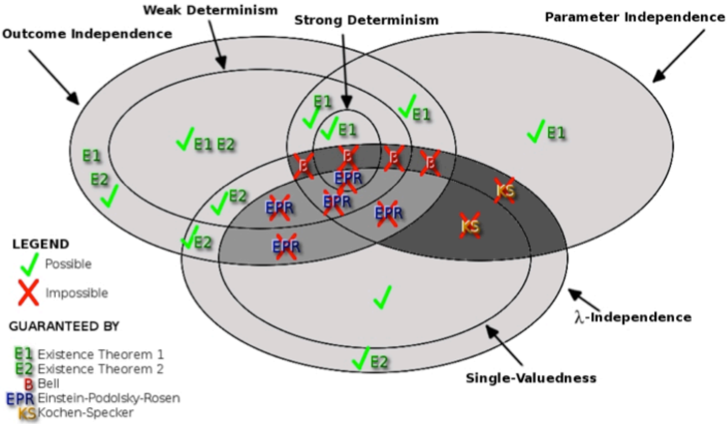
Observational teams

After m rounds of making the experiments $q_1,\ldots,q_n$ we have the data

$$X = \begin{array}{|ccccc|} x_1 & y_1 & \ldots & x_n & y_n \\ \hline a_1^1 & b_1^1 & \ldots & a_n^1 & b_n^1 \\ a_1^2 & b_1^2 & \ldots & a_n^2 & b_n^2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_1^m & b_1^m & \ldots & a_n^m & b_n^m \end{array}$$

Concept arising in quantum foundations

- Determinism
- Hidden variables
- No-signalling
- Outcome independence
- Parameter independence



Picture by Noson Yanofsky in "A Classification of Hidden-Variable Properties", Workshop on Quantum Logic Inspired by Quantum Computation, Indiana, 2009.

Bell's Theorem 1964

- Bell's Theorem in quantum foundations and quantum information theory, the basis of quantum computation, can be seen as the existence of a team, even arising from real physical experiments, violating a dependence logic sentence, which expresses the (falsely) assumed locality, or non-contextuality, of quantum world. (Joint work with Abramsky, Hyttinen and Paolini).
- A logical form of Bell's Theorem in quantum foundations (Hyttinen-Paolini 2014).

Approximate dependence

- Approximate dependence is a much more common phenomenon in science and humanities than full dependence $=\!(x, y)$. Any database of a significant size contains errors for merely human reasons or for reasons of errors in transmission. Any statistical data of medical, biological, social, etc information has exceptions partly because of the nature of the data.
- One rarely if ever encounters absolute dependence of the kind

$$=\!(x, y)$$

in practical examples.

Approximate truth

- “I park the car next to the lamp post depending only on the day of the week, apart from a few exceptions.”,

Example

An employee's salary depends only on the department except for one person.

Employee	Department	Salary
John	I	120 000
Mary	II	130 000
Ann	I	120 000
Paul	I	120 000
Matt	II	130 000
Julia	I	130 000

Table 2

Approximate dependence

Definition

Suppose p is a real number, $0 \leq p \leq 1$. A finite team X is said to satisfy the *approximate dependence atom*

$$=_p(x, y)$$

if there is $Y \subseteq X$, $|Y| \leq p \cdot |X|$, such that the team $X \setminus Y$ satisfies $=(x, y)$. We then write

$$X \models =_p(x, y).$$

Axioms

Definition

The *axioms* of approximate dependence are:

A1 $=_0(xy, x)$ (Reflexivity)

A2 $=_1(x, y)$ (Totality)

The *rules* of approximate dependence are:

A3 If $=_p(x, yv)$, then $=_p(xu, y)$ (Weakening)

A4 If $=_p(x, y)$, then $=_p(xu, yu)$ (Augmentation)

A5 If $=_p(xu, yv)$, then $=_p(ux, yv)$ and $=_p(xu, vy)$ (Permutation)

A6 If $=_p(x, y)$ and $=_q(y, v)$, where $p + q \leq 1$, then $=_{p+q}(x, v)$ (Transitivity)

A7 If $=_p(x, y)$ and $p \leq q \leq 1$, then $=_q(x, y)$ (Monotonicity)

We have the following Completeness Theorem:

Theorem

Suppose Σ is a finite set of approximate dependence atoms. Then

1. $=_p(x, y)$ *follows from Σ by the above axioms and rules*
2. *Every finite team satisfying Σ also satisfies $=_p(x, y)$.*

- We first develop some auxiliary concepts and observations for the proof.
- Let τ be a pair $(\Sigma, =_p(x, y))$, where Σ is a finite set of approximate dependencies. For such τ let Z_τ be the finite set of all variables in $\Sigma \cup \{=_p(x, y)\}$.
- Let C_τ be the smallest set containing Σ and closed under the rules $(A1) - (A6)$ (but not necessarily under (A7)) for variables in Z_τ. Note that C_τ is finite.

Lemma

$\Sigma \vdash =_t(u,v)$ *iff* $\exists r \leq t(=_r(u,v) \in C_\tau)$.

Proof.

The implication from right to left is trivial. For the converse it suffices to show that the set

$$\Sigma' = \{=_t(u,v) : \exists r \leq t(=_r(u,v) \in C_\tau)\}$$

is closed under (A1)-(A7). □

Definition

Suppose $\tau = (\Sigma, =_p(x, y))$. For any variable y let

$$d_\tau(y) = \min\{r \in [0,1] : =_r(x,y) \in C_\tau\}.$$

This definition makes sense because there are only finitely many $=_r(u, v)$ in C_τ. Note that $d_\tau(x) = 0$ by axiom (A1). By Lemma 24,

$$d_\tau(y) = \min\{r \in [0,1] : \Sigma \vdash =_r(x,y)\}.$$

Lemma

If $\Sigma \vdash =_p(u, v)$, *then* $d_\tau(v) - d_\tau(u) \leq p$.

- For a given Σ there are only finitely many numbers $d_\tau(u)$, $u \in Z_\tau$, because C_τ is finite. Let A_τ consist of p and the set of $d_\tau(u)$ such that $u \in Z_\tau$. Let $n = 1 + \max\{\lceil 2/(a-b)\rceil : a, b \in A - \tau, a \neq b\}$. We define a team X_τ of size n as follows:

$$X_\tau = \{s_0, \ldots, s_n\},$$

where for $\frac{m}{n} \leq d_\tau(u) < \frac{m+1}{n}$ we let

$$s_i(u) = \begin{cases} i, & \text{if } i \leq m \\ m, & \text{if } i > m \end{cases}$$

	x	$\dots$	u	$\dots$
s_0	0	$\dots$	0	$\dots$
s_1	0	$\dots$	1	$\dots$
s_2	0	$\dots$	2	$\dots$
$\vdots$				
s_m	0	$\dots$	m	$\dots$
$\vdots$				
s_{n-1}	0	$\dots$	m	$\dots$

Figure: The team X_τ

Lemma

Suppose $X_\tau \models =_p(x,y)$*. Then* $\Sigma \vdash =_p(x,y)$*.*

Proof.

Suppose $X_\tau \models =_p(x,y)$ but $\Sigma \nvdash =_p(x,y)$. Now $d_\tau(y) > p$. Let $\frac{m}{n} \leq d_\tau(y) < \frac{m+1}{n}$. One has to take all the assignments s_i, $i \leq m-1$, away from X_τ in order for the remainder to satisfy $=(x,y)$. Hence $p \cdot n \geq m$ i.e. $p \geq \frac{m}{n}$. But we have chosen n so that $1/n < d_\tau(y) - p$. Hence

$$p < d_\tau(y) - \frac{1}{n} \leq \frac{m+1}{n} - \frac{1}{n} = \frac{m}{n},$$

a contradiction. □

Lemma

Suppose $\Sigma \vdash =_q(u,v)$. *Then* $X_\tau \models =_q(u,v)$.

Proof.

- We know already $d_\tau(v) - d_\tau(u) \leq q$. If $d_\tau(v) \leq d_\tau(u)$, then $X_\tau \models =(u,v)$, and hence all the more $X_\tau \models =_q(u,v)$. Let us therefore assume $d_\tau(v) > d_\tau(u)$. Since $2/n < d_\tau(v) - d_\tau(u)$, there are m and k such that

$$\frac{m}{n} \leq d_\tau(u) < \frac{m+1}{n} < \frac{k}{n} \leq d_\tau(v) < \frac{k+1}{n}.$$

- In order to satisfy $=(x,y)$ one has to delete $k-m$ assignments from X_τ. But this is fine, as $qn \geq (d_\tau(v) - d_\tau(u))n \geq k - d_\tau(u)n \geq k - m$.

□

- The canonical example of a team in dependence logic is the set of plays where a player is using a fixed strategy. Such a team satisfies certain dependence atoms reflecting commitments the player has made concerning information he or she is using.
- If such dependence atoms hold only approximatively, the player is allowed to make a small number of deviations from his or her commitments.

Punchline

- The emergent logic of dependence and independence provides a common mathematical basis for fundamental concepts in biology, social science, physics, mathematics and computer science.
- We can find fundamental principles governing this logic.
- Algorithmic results show—as can be expected—that dependence logic has higher complexity than ordinary first order (propositional, modal) logic.
- Important parts can be completely axiomatized, other parts are manifestly beyond the reach of axiomatization.

Thank you!