Dependence and Axiomatizations

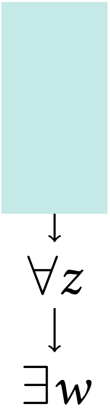
Logic seminar in Göteborg

Fredrik Engström

September 27, 2013

Intro

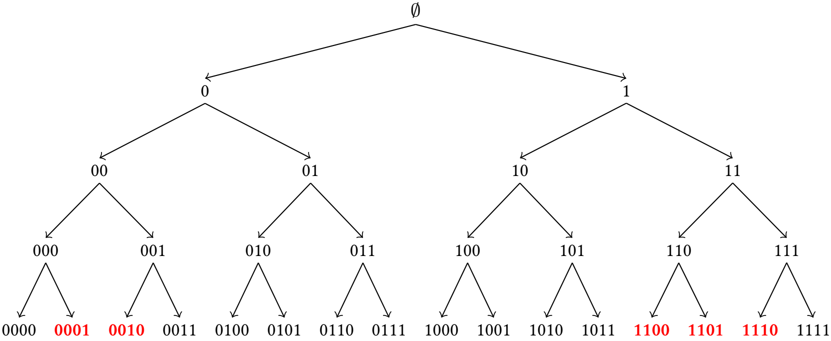
$$\forall x \exists y \forall z \exists w Rxyzw$$

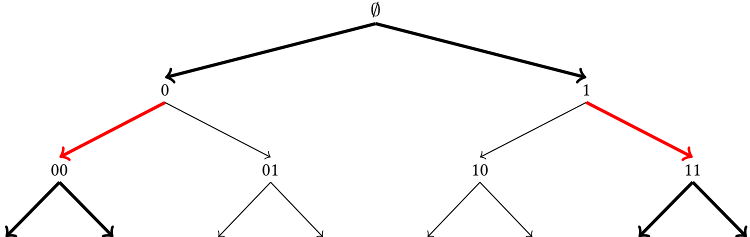


$$\begin{pmatrix} \forall x \exists y \\ \forall z \exists w \end{pmatrix} Rxyzw$$

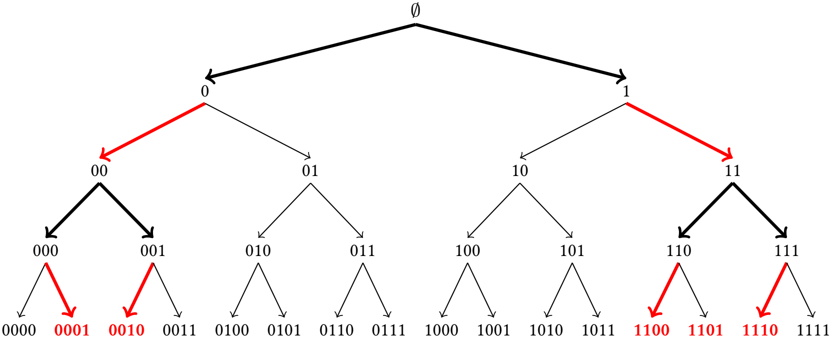
$$\begin{matrix} \forall z & \forall x \\ \downarrow & \downarrow \\ \exists w & \exists y \end{matrix}$$

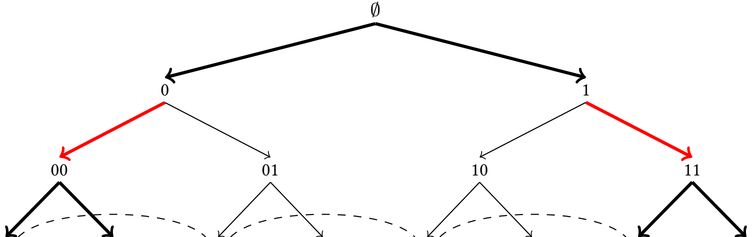
Domain $\{0, 1\}$. $\forall x \exists y \forall z \exists w Rxyzw$





Domain $\{ 0, 1 \}$. $\begin{pmatrix} \forall x \exists y \\ \forall z \exists w \end{pmatrix} Rxyzw$





x	y	z	w
0	0	0	1
0	0	1	0
1	1	0	0
1	1	1	0

$\not\models \; =\!(z, w)$

x	y	z	w
0	0	0	1
0	0	1	0
1	1	0	1
1	1	1	0

$\models \; =\!(z, w)$

$$\begin{pmatrix} \forall x \exists y \\ \forall z \exists w \end{pmatrix} Rxyzw \equiv \forall x \exists y \forall z \exists w (=\!(z, w) \wedge Rxyzw)$$

Definition

X a **team** = set of assignments.

Dependence and axiomatizations

Normal forms I

A database in normal-form is less sensitive to modification anomalies, by reducing the number of dependencies in the database.

Name	SSN	Course
Svensson	910101−0101	Logic
Svensson	910101−0101	Philosophy
Svensson	920202−0202	Logic
Olsson	930303−0303	Philosophy

Here SSN $\rightarrow$ Name.

Name	SSN
Svensson	910101−0101
Svensson	920202−0202
Olsson	930303−0303

SSN	Course
910101−0101	Logic
910101−0101	Philosophy
920202−0202	Logic
930303−0303	Philosophy

2NF, 3NF, and BCNF (Boyce−Codd normal form) puts restrictions on functional dependencies.

Dependencies

Equality generating dependencies:

- Functional dependence $\bar{x} \to \bar{y}$: The y's are functional determined by the x's.

$$\forall s, s' \in X\big(\bigwedge s(x_i) = s'(x_i) \to \bigwedge s(y_i) = s'(y_i)\big)$$

Tuple generating dependencies:

- Multivalued dependence $\bar{x} \twoheadrightarrow \bar{y}$. The set of possible values of $\bar{y}$ is determined $\bar{x}$. $X \vDash \bar{x} \twoheadrightarrow \bar{y}$ iff

  for all $s, s' \in X$ if $s(\bar{x}) = s'(x)$ then there is $s_0 \in X, s_0(\bar{x}, \bar{y}) = s(\bar{x}, \bar{y})$ and $s_0(\bar{z}) = s'(\bar{z})$, where $\bar{z}$ complement of $\bar{x} \cup \bar{y}$.

$$\forall s, s' \in X\big(\bigwedge s(x_i) = s'(x_i) \to \exists s_0 \bigwedge t_i = t_i'\big)$$

- Embedded multivalued dependence $\bar{x} \twoheadrightarrow \bar{y}|\bar{z}$. The projection to $\bar{x}, \bar{y}, \bar{z}$ satisfies $\bar{x} \twoheadrightarrow \bar{y}$.
- Template dependence.

Normal forms II

Student	Course	Lecturer
Svensson	Logic	Lindström
Svensson	Philosophy	Lindström
Svensson	Philosophy	Bennet
Olsson	Philosophy	Lindström
Olsson	Philosophy	Bennet

Not Course $\to$ Lecturer, but Course $\twoheadrightarrow$ Lecturer.

Student	Course
Svensson	Logic
Svensson	Philosophy
Olsson	Philosophy

Course	Lecturer
Logic	Lindström
Philosophy	Lindström
Philosophy	Bennet

- 4NF reduces multivalued dependencies. 5NF join dependencies.

Remarks

$D \cup \{ \phi \}$ is a (finite) set of dependence relations.

$$D \vDash \phi \text{ if } \forall X (X \vDash D \Rightarrow X \vDash \phi).$$

- Any first-order definable dependence has an r.e. entailment relation.
- In some (most) cases we have decidability of the entailment relation.

Functional dependence

Definition

$X \vDash \bar{x} \rightarrow \bar{y}$ iff for all $s, s' \in X$ if $s(\bar{x}) = s'(\bar{x})$ then $s(\bar{y}) = s'(\bar{y})$.

Theorem (Armstrong, 1974)

$D \vDash \phi$ iff ϕ is derivable from D using the rules:

- **Reflexivity:** If $\bar{y} \subseteq \bar{x}$ then $\bar{x} \rightarrow \bar{y}$.
- **Augmentation:** If $\bar{x} \rightarrow \bar{y}$ then $\bar{x}, z \rightarrow \bar{y}, z$.
- **Transitivity:** If $\bar{x} \rightarrow \bar{y}$ and $\bar{y} \rightarrow \bar{z}$ then $\bar{x} \rightarrow \bar{z}$.

$D \vDash \phi$ is decidable in time $O(n)$.

Multivalued dependence

Fix a set U of variables.

Theorem (Beeri, Fagin, Howard, 1977)

$D \vDash_U \phi$ iff ϕ is derivable from D with the following inference rules:

- **Complementation:** If $\bar{x} \cup \bar{y} \cup \bar{z} = U$, $\bar{y} \cap \bar{z} \subseteq \bar{x}$, and $\bar{x} \twoheadrightarrow \bar{y}$ then $\bar{x} \twoheadrightarrow \bar{z}$
- **Reflexivity:** If $\bar{y} \subseteq \bar{x}$ then $\bar{x} \twoheadrightarrow \bar{y}$.
- **Augmentation:** If $\bar{z} \subseteq \bar{w}$ and $\bar{x} \twoheadrightarrow \bar{y}$ then $\bar{x}, \bar{w} \twoheadrightarrow \bar{y}, \bar{z}$.
- **Transitivity:** If $\bar{x} \twoheadrightarrow \bar{y}$ and $\bar{y} \twoheadrightarrow \bar{z}$ then $\bar{x} \twoheadrightarrow \bar{z} \setminus \bar{y}$.

Solvable in time $O(n \log n)$.

Template dependencies

$X \models \bar{x} \twoheadrightarrow \bar{y}$ iff for all $s, s' \in X$ if $s(\bar{x}) = s'(x)$ then there is $s_0 \in X$, $s_0(\bar{x}, \bar{y}) = s(\bar{x}, \bar{y})$ and $s_0(\bar{z}) = s'(\bar{z})$, where $\bar{z}$ are the variables not in $\bar{x}$ or in $\bar{y}$.

Tuple-generating dependencies are all of similar form.

x	y	z	w
a	b	c	d
a	b	c'	d'
a	b	c	d'

$\equiv xy \twoheadrightarrow z$

Let R be a relation and s one assignment. $X \models (R, s)$ iff for every embedding of R into X there is an extension of the embedding also embedding s in X.

Chasing

Chase step

Given team X and TD (R, s) and an embedding of R into X, if that embedding can't be extended to include s, add a new row to X.

The chase of X according to a set t of TDs is defined similar. The result of a chase is denoted $\text{chase}_t(X)$.

- For template dependencies and multivalued dependencies this procedure terminates. Not so for embedded multivalued dependence.
- Can be used to show that the implication problem for template dependencies is decidable: $\Gamma \models (R, s)$ iff s 'appears' in $\text{chase}_\Gamma(R)$.
- Can also be used to design a (sound and complete) deduction system for template deductions.
- Also an r.e. algorithm for deciding implication for EMVD.

Embedded multivalued dependencies

Theorem (Herrmann, 1995)
The implication problem for embedded multivalued dependencies (EMVD) is undecidable.

Theorem (Sagiv and Walecka, 1982)
There is no sound and complete deduction system for EMVD with a bound on the number of premisses in the rules.

Theorem (Hannula and Kontinen, 2013)
There is a sound and complete axiomatization of EMVD together with inclusion dependencies.

Open question: Is there a sound and complete axiomatization for EMVD? (No system bounded in U).

Dependence logic

Generalized quantifiers

- **Generalized quantifier**: Q a class of structures (one unary relation).
- $M, s \models Qx\,\phi$ iff $(M, \phi^{M,s}) \in Q$.
- $Q_M = \{ R \mid (M, R) \in Q \}$.

- $\mathrm{FO}(Q)$ is FOL extended with expressions $Qx\,\phi$.
- $\exists$ and $\forall$ can be interpreted as generalized quantifiers.
- We will **only consider (non-trivial) monotone increasing quantifiers**: If $A \subseteq B$ and $A \in Q_M$ then $B \in Q_M$.
- $\check{Q} = \{ (M, R^c) \mid (M, R) \notin Q \}$. $\neg Qx \neg\phi \equiv \check{Q}x\,\phi$.

Dependence Logic with Q

- $D(Q)$ is $\phi ::= \gamma \mid \phi \wedge \phi \mid \phi \vee \phi \mid \exists x \phi \mid \forall x \phi \mid Qx \phi$,
  γ an $\mathrm{FO}(Q)$ formula or (functional) dependence atom: $\bar{x} \to y$.
- $M \vDash \sigma$ iff $M, \{\emptyset\} \vDash \sigma$.

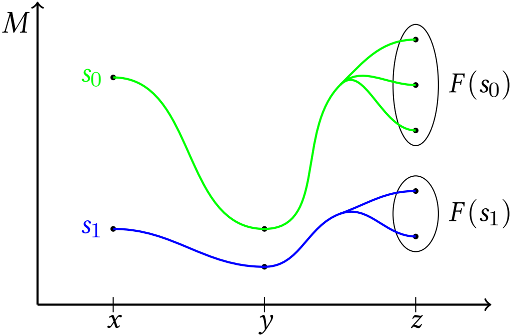
- $M, X \vDash \gamma$ if for all $s \in X$: $M, s \vDash \gamma$,
  γ an $\mathrm{FO}(Q)$ formula.
- $M, X \vDash \phi \wedge \psi$ iff $M, X \vDash \phi$ and $M, X \vDash \psi$.
- $M, X \vDash \phi \vee \psi$ iff there are $Y \cup Z = X$ such that $M, Y \vDash \phi$ and $M, Z \vDash \psi$.

Quantifiers in dependence logic

- $M, X \models Qx\,\phi$ iff there is $F : X \to Q_M$ such that $M, X[F/x] \models \phi$.

$$X[F/x] = \{ s[a/x] \mid s \in X, a \in F(s) \}.$$

Example: $M, \{s_0, s_1\} \models Qz\,Rxyz$



- Note: using set-valued Fs corresponds to **non-deterministic** strategies.

Properties of Dependence logic

- $M, \emptyset \models \phi$
- **Downwards closure:** If $Y \subseteq X$ and $M, X \models \phi$ then $M, Y \models \phi$.
- $D(Q) \equiv \mathrm{ESO}(Q)$ (E / Kontinen)

Theorem: Normal form for $D(Q, \breve{Q})$

Every $D(Q, \breve{Q})$ formula is equivalent to one of the form:

$$\mathcal{H}^1 x_1 \ldots \mathcal{H}^m x_m \exists y_1 \ldots \exists y_n \Big(\bigwedge_{1 \leq i \leq n} (\overline{x}^i \to y_i) \wedge \theta \Big),$$

where $\mathcal{H}^i$ is either Q, $\breve{Q}$ or $\forall$, and θ is a quantifier-free FO formula.

Example

$$\exists a \forall x \exists y \forall z \exists w \big((z \to w) \land (x = z \leftrightarrow y = w) \land y \neq a\big)$$

$$\exists a \exists f, g \forall x, z \big((x = z \leftrightarrow f(x) = g(z)) \land f(x) \neq a\big)$$

$$\exists a \exists f \forall x, z \big((f(x) = f(z) \to x = z) \land f(x) \neq a\big)$$

Digression: Multivalued dependence and quantifiers

Remember:

$$\begin{pmatrix} \forall x \exists y \\ \forall z \exists w \end{pmatrix} Rxyzw \equiv \forall x \exists y \forall z \exists w (z \to w \wedge Rxyzw)$$

Not true for generalized quantifiers.

However:

$$\begin{pmatrix} Q_1 x \\ Q_2 y \end{pmatrix} Rxyz \equiv Q_1 x Q_2 y (z \twoheadrightarrow y \wedge Rxyz)$$

$$\begin{aligned} \exists w, w' \big(\quad & (z \to w) \wedge (z \to w') \wedge \\ & Qx \exists y (y = w \wedge (\bar{z}, x \to y) \wedge \\ & Qx' \exists y' (y' = w' \wedge (z, x' \to y') \wedge \\ & \forall u \exists v ((z, u \to v) \wedge (x = u \to v = w) \wedge \\ & \forall u' \exists v' ((z, u' \to v') \wedge (x' = u' \to v' = w') \wedge \\ & ((v = w \wedge v' = w') \to \phi(u, u', z)))))\big) \end{aligned}$$

Axiomatizations

Dependence logic

- Dependence logic has the same strength as ESO.
- The relation $\Gamma \vDash \phi$ is not r.e.
- Open question (to me at least): On what level of complexity does the non-r.e. appear. Compare:

  Theorem (Kontinen, 2013)
  Model checking for the disjunction of two dependence atoms is not NP-complete, but the disjunction of three dependence atoms has a NP-complete model checking.

- Restricting to ϕ's without dependence atoms gives an r.e. entailment relation.

General idea for axiomatizations

Idea:

- Construct a natural deduction system in which the normal form can be derived.
- Allow dependencies in normal forms to be replaced by **finite approximations**.
- Show that in enough models (recursively saturated) the set of finite approximations is equivalent to the original sentence.

Axiomatizing $D(Q, \check{Q})$ I: General rules

First: Some rules sound for **any interpretation** of Q (monotone increasing).

- Standard rules for $\mathrm{FO}(Q, \check{Q})$ formulas.
- Standard rules for conjunction, existential quantifier, and universal quantifier.
- Commutativity, associativity and monotonicity of disjunction.
- Monotonicity, extending scope, and renaming of bound variables for Q and $\check{Q}$.
- Duality of $\check{Q}$ with respect to $\mathrm{FO}(Q, \check{Q})$ formulas.

Axiomatizing $D(Q, \check{Q})$ II: Dependence related rules

- Unnesting:

$$\frac{(t_1, ..., t_{n-1} \to t_n)}{\exists z((t_1, ..., z, ..., t_{n-1} \to t_n) \wedge z = t_i)}$$

where z is a new variable.

- Dependence distribution:

$$\frac{\exists y_1 \ldots \exists y_n(\bigwedge_{1 \leq j \leq n}(\bar{z}^j \to y_j) \wedge \phi) \vee \exists y_{n+1} \ldots \exists y_m(\bigwedge_{n+1 \leq j \leq m}(\bar{z}^j \to y_j) \wedge \psi}{\exists y_1 \ldots \exists y_m(\bigwedge_{1 \leq j \leq m}(\bar{z}^j \to y_j) \wedge (\phi \vee \psi))}$$

where ϕ and ψ are quantifier free FO formulas.

- Dependence introduction:

$$\frac{\exists x \mathcal{H} y \phi}{\mathcal{H} y \exists x((\bar{z} \to x) \wedge \phi)}$$

where $\bar{z}$ lists the variables in $\mathrm{FV}(\phi) - \{x, y\}$ and $\mathcal{H} \in \{\forall, Q, \check{Q}\}$.

Approximations

Suppose σ is in normal form:

$$\mathcal{H}^1 x_1 \dots \mathcal{H}^m x_m \exists y_1 \dots \exists y_n \Big(\bigwedge_{1 \le i \le n} (\bar{x}^i \to y_i) \wedge \theta(\bar{x}, \bar{y}) \Big).$$

Let $A^k \sigma$ be

$$\forall \bar{x}_1 \exists \bar{y}_1 \dots \forall \bar{x}_k \exists \bar{y}_k \Big(\bigwedge_{1 \le j \le k} R(\bar{x}_j) \to \bigwedge_{1 \le j \le k} \theta(\bar{x}_j, \bar{y}_j) \wedge \bigwedge_{\substack{1 \le i \le n \\ 1 \le j, j' \le k}} (\bar{x}^i_j = \bar{x}^i_{j'} \to y_{i,j} = y_{i,j'}) \Big)$$

Let $B\sigma$ be

$$\mathcal{H}^1 x_1 \dots \mathcal{H}^m x_m R(x_1, \dots, x_m).$$

Axiomatizing $D(Q, \breve{Q})$ III: The approximation rule

$$\dfrac{\sigma \qquad \begin{matrix} [B\,\sigma] & & [A^k\,\sigma] \\ \ddots & & \iddots \\ & \psi & \end{matrix}}{\psi}\ (\text{Approx})$$

where σ is a sentence in normal form, and R does not appear in ψ nor in any uncancelled assumptions in the derivation of ψ, except for $B\,\sigma$ and $A^k\,\sigma$.

Completeness for weak semantics

Let $\Gamma \vDash_w \phi$ mean that $\Gamma \vDash \phi$ for any monotone increasing (non-trivial) interpretation of Q (and $\check{Q}$ is interpreted as the dual of the interpretation of Q).

Theorem

This system is sound and complete wrt $\Gamma \vDash_w \phi$ where ϕ is $\mathrm{FO}(Q, \check{Q})$.

Uncountably many

- $\mathrm{FO}(Q_1)$ is axiomatizable, where Q_1 is "there exist uncountably many ...". (Kiesler)
- Add Keisler's rules for Q_1.

Define the **Skolem translation** $S\sigma$ of σ in normal form to be:

$$\mathcal{H}^1 x_1 \dots \mathcal{H}^m x_m \theta(f_i(\bar{x}^i)/y_i).$$

- Replace the approximation rule with the following rule

$$\dfrac{\sigma \qquad \begin{matrix}[S\sigma] \\ \vdots \\ \psi\end{matrix}}{\psi} \text{ (Skolem)}$$

Theorem
This system is sound and complete wrt $\Gamma \vDash \phi$ where ϕ is $\mathrm{FO}(Q_1, \check{Q}_1)$.

Further work

- Is there a deduction system for EMVD?
- Which (weak) fragments of dependence logic has en r.e. entailment relation?
- Are there axiomatizations of first-order consequences of dependence logic not based on normal forms?
- Is it possible to axiomatize $D(Q)$ within the given language?
- Is it possible to axiomatize $D(Q_1)$ without using the Skolem normal form.

Thank you for your attention.

W. W. Armstrong. Dependency structures of database relationships. In Proceedings of the ifip Congress. 1974.

Fredrik Engström. Generalized quantifiers in dependence logic. Journal of Logic, Language and Information, 21:299–324, 2012. ISSN 0925-8531.

Fredrik Engström and Juha Kontinen. Characterizing quantifier extensions of dependence logic. Journal of Symbolic Logic, 78(1):307–316, 2013.

Fredrik Engström, Juha Kontinen, and Jouko Väänänen. Dependence logic with generalized quantifiers: Axiomatizations. In Logic, Language, Information, and Computation, 2013.

Miika Hannula and Juha Kontinen. A finite axiomatization of conditional independence and inclusion dependencies. arXiv preprint arXiv:1309.4927, 2013.

Leon Henkin. Some remarks on infinitely long formulas. In Infinitistic Methods (Proc. Sympos. Foundations of Math., Warsaw, 1959), pages 167–183. Pergamon, Oxford, 1961.

Christian Herrmann. On the undecidability of implications between embedded multivalued database dependencies. Information and Computation, 122(2):221–235, 1995.

Wilfrid Hodges. Compositional semantics for a language of imperfect information. Logic Journal of IGPL, 5(4):539–563, 1997.

H Jerome Keisler. Logic with the quantifier “there exist uncountably many”. Annals of Mathematical Logic, 1(1):1–93, 1970.

Jarmo Kontinen. Coherence and computational complexity of quantifier-free dependence logic formulas. Studia Logica, pages 1–25, 2013.

Juha Kontinen and Jouko Väänänen. Axiomatizing first-order consequences in dependence logic. Annals of Pure and Applied Logic, 164(11):1101–1117, 2013.

Yehoshua Sagiv and Scott F Walecka. Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies. Journal of the ACM (JACM), 29(1):103–117, 1982.

Jouko Väänänen. Dependence logic, volume 70 of London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 2007. ISBN 978-0-521-70015-3; 0-521-70015-9. A new approach to independence friendly logic.