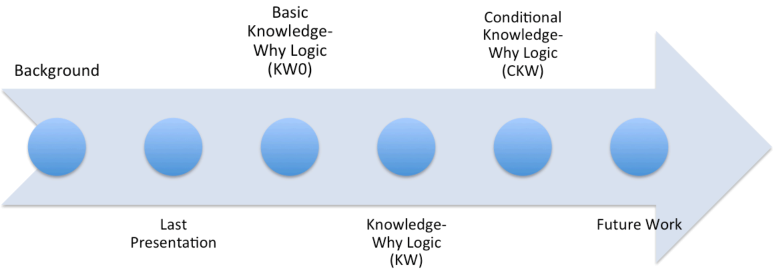
The Logic of Knowledge-Why (2)

Yanjing Wang, Chao Xu

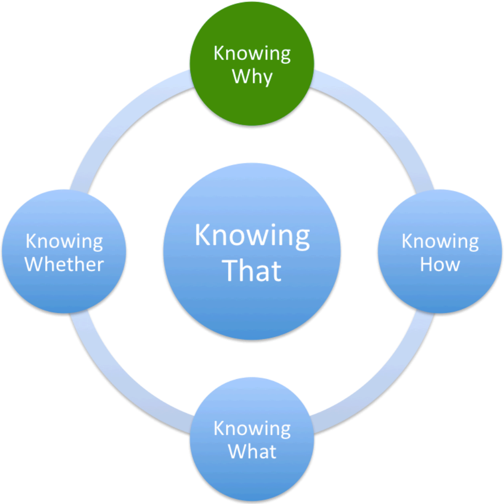
Chao Xu

Department of Philosophy, Peking University

May 5th, 2015


Background
Basic Knowledge-Why Logic (KW0)
Conditional Knowledge-Why Logic (CKW)
Last Presentation
Knowledge-Why Logic (KW)
Future Work

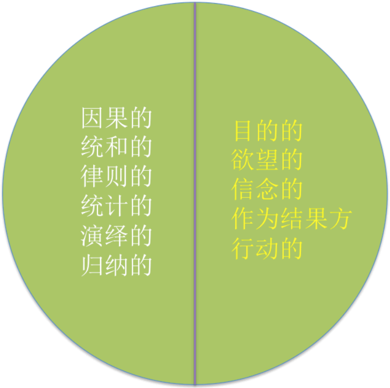
§1 Background



When I say "I know why he comes late.", I must know an explanation, a cause or a reason for why he comes late, not only the event that he comes late. In the other words, I must know the answer to the question "Why did he come late?"

- Knowledge-Why is a kind of knowledge such as 'I know why φ'.
- Since explanations can often be thought of as answers to why-questions, 'I know why φ' means 'I know the explanation of φ'. To study the knowledge why, we first need to know what the explanation of φ is.
- The main study about the explanation is in the field of scientific explanation, and thus we also study the explanation on the basis of the theory of scientific explanation.

The types of explanations



It is the foundation of our study that explanation is an argument.

§2 Last Presentation

- Explanation and Scientific Explanation
- The Logic of Justification
- The Logic of Knowledge-Why

Language:(single-agent)

$$\varphi = \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathcal{K}\varphi \mid \mathcal{K}_{why}\varphi$$

- A1 Classical Propositional Axioms
- A2 $\mathcal{K}(\varphi \to \psi) \to (\mathcal{K}\varphi \to \mathcal{K}\psi)$
- A3 $\mathcal{K}_{why}(\varphi \to \psi) \to (\mathcal{K}_{why}\varphi \to \mathcal{K}_{why}\psi)$
- A4 $\mathcal{K}_{why}\varphi \to \mathcal{K}\varphi$
- A5 $\mathcal{K}_{why}\varphi \to \mathcal{K}\mathcal{K}_{why}\varphi$
- A6 $\mathcal{K}\varphi \to \mathcal{K}\mathcal{K}\varphi$
- R1 Modus Ponens
- R2 $\vdash \varphi \quad \Rightarrow \quad \vdash \mathcal{K}\varphi$

Basic Epistemic Semantics

Definition (KW-model: $\mathcal{M} = (W, E, R, \mathcal{E}, V)$)

W: The set of possible worlds

E: The set of explanations

R: The accessible relation between the worlds in W, R is transitive.

$\mathcal{E}$: $\mathcal{E}(t, \varphi) \subseteq W$ specifies the set of possible worlds where t is considered admissible explanation for φ.
An admissible explanation function $\mathcal{E}$ must satisfy the conditions: $\forall r, s$, If $w \in \mathcal{E}(r, \varphi \to \psi) \cap \mathcal{E}(s, \varphi)$, then there exists t such that $w \in \mathcal{E}(t, \psi)$ and $v \in \mathcal{E}(t, \psi)$ for all v such that wRv and $v \in \mathcal{E}(r, \varphi \to \psi) \cap \mathcal{E}(s, \varphi)$

V: Atom $\to \mathcal{P}(W)$

Now, we can define the satisfiable relation $\Vdash$:

- $w \Vdash \top$
- $w \Vdash p$ iff $w \in V(p)$
- $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$
- $w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$
- $w \Vdash \mathcal{K}\varphi$ iff for each v such that wRv, $v \Vdash \varphi$
- $w \Vdash \mathcal{K}_{why}\varphi$ iff (1)$\exists t, w \in \mathcal{E}(t, \varphi), \forall v, \ wRv, v \in \mathcal{E}(t, \varphi)$ and (2) $\forall v, \ wRv, v \Vdash \varphi$

Soundness and Completeness

Theorem

KW is sound and complete for the class of all KW-models.

Completeness:

To establish completeness, we use standard canonical model construction. The canonical model $\mathcal{M}^c = (W^c, E^c, R^c, \mathcal{E}^c, V^c)$ for KW is defined as follows:

Let *Form* be the set of all formulas. Define $\Sigma = \{f \mid f : Form \rightsquigarrow E^c, f \text{ is a partial function}\}$. For each $f \in \Sigma$, f satisfies the condition as follows: If $f(\varphi \to \psi) = r$ and $f(\varphi) = s$, then there exists t such that $f(\psi) = t$.

- $W^c = \{\langle\Gamma, f\rangle \mid \langle\Gamma, f\rangle \in$MCS$\times\Sigma$, If $\mathcal{K}_{why}\varphi \in \Gamma$, then there exists t such that $f(\varphi) = t\}$, MCS is the set of all maximal consistent sets in KW. Following an established tradition, we denote elements of W^c as $\langle\Gamma, f\rangle, \langle\Delta, g\rangle$, and so forth;
- $\langle\Gamma, f\rangle R^c \langle\Delta, g\rangle$ iff for all formulas such as $\mathcal{K}_{why}\varphi \in \Gamma$, $f(\varphi) = g(\varphi)$ and $\Gamma^{\#} \subseteq \Delta$, where
  $\Gamma^{\#} = \{\varphi \mid \mathcal{K}_{why}\varphi \in \Gamma\} \cup \{\mathcal{K}_{why}\varphi \mid \mathcal{K}_{why}\varphi \in \Gamma\} \cup \{\varphi \mid \mathcal{K}\varphi \in \Gamma\}$
- $\mathcal{E}^c(t, \varphi) = \{\langle\Gamma, f\rangle \mid \mathcal{K}_{why}\varphi \in \Gamma, f(\varphi) = t\}$
- $V^c(p) = \{\langle\Gamma, f\rangle \mid p \in \Gamma\}$

§3 Basic Knowledge-Why Logic

Basic Justification Logic J_0 (Artemov, 2008)

- Classical propositional axioms and rule Modus Ponens
- $s : (F \to G) \to (t : F \to (s \cdot t) : G)$ Application
- $s : F \to (s + t) : F, t : F \to (s + t) : F$ Sum

Factivity Axiom ($t : F \to F$) is not required in basic Justification Logic systems.

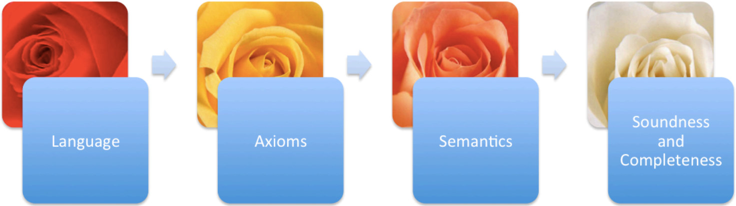
Language(single-agent)

$$\varphi = \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathcal{K}_{why}\varphi$$

- A1 Classical Propositional Axioms
- A2 $\mathcal{K}_{why}(\varphi \to \psi) \to (\mathcal{K}_{why}\varphi \to \mathcal{K}_{why}\psi)$
- A3 $\mathcal{K}_{why}\varphi \to \varphi$
- R1 Modus ponens

KW_0 could be viewed as the correspondence epistemic modal logic of Justification Logic J_0T. ($J_0T = J_0 + T$).

§4 Knowledge-Why Logic (abbr. KW)



Language(single-agent): $\varphi = \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathcal{K}\varphi \mid \mathcal{K}_{why}\varphi$

- A1 Classical Propositional Axioms
- A2 $\mathcal{K}(\varphi \to \psi) \to (\mathcal{K}\varphi \to \mathcal{K}\psi)$
- A3 $\mathcal{K}_{why}(\varphi \to \psi) \to (\mathcal{K}_{why}\varphi \to \mathcal{K}_{why}\psi)$
- A4 $\mathcal{K}\varphi \to \varphi$
- A5 $\mathcal{K}_{why}\varphi \to \mathcal{K}\varphi$
- A6 $\mathcal{K}\varphi \to \mathcal{K}_{why}\mathcal{K}\varphi$
- A7 $\neg\mathcal{K}\varphi \to \mathcal{K}_{why}\neg\mathcal{K}\varphi$
- A8 $\mathcal{K}_{why}\varphi \to \mathcal{K}\mathcal{K}_{why}\varphi$
- A9 $\neg\mathcal{K}_{why}\varphi \to \mathcal{K}\neg\mathcal{K}_{why}\varphi$
- A10 $\mathcal{K}_{why}(\varphi \wedge \psi) \leftrightarrow \mathcal{K}_{why}\varphi \wedge \mathcal{K}_{why}\psi$

Rules:

R1: Modus Ponens

R2: $\vdash \varphi \quad \Rightarrow \quad \vdash \mathcal{K}\varphi$

Propositions

(1) $\mathcal{K}\varphi \rightarrow \mathcal{K}\mathcal{K}\varphi$

(2) $\neg\mathcal{K}\varphi \rightarrow \mathcal{K}\neg\mathcal{K}\varphi$

(3) $\mathcal{K}_{why}\varphi \rightarrow \varphi$

Basic Epistemic Semantics $\mathcal{M} = (W, E, R, \mathcal{E}, V)$

W: The set of possible worlds.

E: The set of explanations. E satisfy the closure condition: If $s, t \in E$, then $s \cdot t, s + t \in E$. $'e'$, a special element in E, is called 'introspective explanation'. A possible motivation for considering such an explanation could be that $'e'$ may be well regarded as the explanation of the formulas such as $\mathcal{K}\varphi$ or $\neg\mathcal{K}\varphi$.

R: The accessible relation between the worlds in W. R is an equivalence relation. (i.e. R is reflexive, transitive and symmetry).

$\mathcal{E}$: Admissible explanation function such that $\mathcal{E}(t, \varphi) \subseteq W$ for any formula $t \in E$ and formula φ. $\mathcal{E}(t, \varphi) \subseteq W$ specifies the set of possible worlds where t is considered as admissible explanation for φ.

An admissible explanation function $\mathcal{E}$ must satisfy the conditions as follows:

(1) $\mathcal{E}(s, \varphi \to \psi) \cap \mathcal{E}(t, \varphi) \subseteq \mathcal{E}(s \cdot t, \psi)$. This condition states that whenever s is an admissible explanation for $\varphi \to \psi$ and t is an admissible explanation for φ, $s \cdot t$ is an admissible explanation for ψ.

(2) $\mathcal{E}(t, \varphi \wedge \psi) \subseteq \mathcal{E}(t, \varphi) \cap \mathcal{E}(t, \psi)$.

(3) $\mathcal{E}(s, \varphi) \cap \mathcal{E}(t, \psi) \subseteq \mathcal{E}(s + t, \varphi \wedge \psi)$

(4) $\mathcal{E}(e, \varphi) = W$, φ is the formula such as $\mathcal{K}\psi$ or $\neg\mathcal{K}\psi$.

[V:] The map from the set of atomic formulas to the power set of the set of possible worlds.

Now, we can define the satisfiable relation $\Vdash$:

- $w \Vdash \top$
- $w \Vdash p$ iff $w \in V(p)$
- $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$
- $w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$
- $w \Vdash \mathcal{K}\varphi$ iff for each v such that wRv, $v \Vdash \varphi$
- $w \Vdash \mathcal{K}_{why}\varphi$ iff (1)$\exists t, w \in \mathcal{E}(t, \varphi), \forall v,\ wRv, v \in \mathcal{E}(t, \varphi)$ and (2) $\forall v,\ wRv, v \Vdash \varphi$

Theorem

KW is sound and complete for the class of all KW-models.

A6: $\mathcal{K}\varphi \to \mathcal{K}_{why}\mathcal{K}\varphi$

Proof.

Suppose $w \Vdash \mathcal{K}\varphi$. Since R is transitive, then we have $v \Vdash \mathcal{K}\varphi$ for all possible worlds v such that wRv. By the condition(4) of $\mathcal{E}$, for each v such that wRv, there exists $e \in E$ such that $v \in \mathcal{E}(e, \mathcal{K}\varphi)$. Hence we conclude $w \Vdash \mathcal{K}_{why}\mathcal{K}\varphi$. □

A10: $\mathcal{K}_{why}(\varphi \wedge \psi) \leftrightarrow \mathcal{K}_{why}\varphi \wedge \mathcal{K}_{why}\psi$

Proof.

Suppose $w \Vdash \mathcal{K}_{why}(\varphi \wedge \psi)$. By the definition of $\Vdash$, we have that for all v such that wRv: (1)$v \Vdash \varphi \wedge \psi$. (2)there exists t such that $v \in \mathcal{E}(t, \varphi \wedge \psi)$. By the condition(2) of $\mathcal{E}$, we have $v \in \mathcal{E}(t, \varphi) \cap \mathcal{E}(t, \psi)$. Then we have $v \in \mathcal{E}(t, \varphi)$ and $v \in \mathcal{E}(t, \psi)$. We also have $v \Vdash \varphi$ and $v \Vdash \psi$ because of (1). It follows that $w \Vdash \mathcal{K}_{why}\varphi$ and $w \Vdash \mathcal{K}_{why}\psi$. Similarly, suppose $w \Vdash \varphi$ and $w \Vdash \psi$, we can prove $w \Vdash \varphi \wedge \psi$ by the condition(3) of $\mathcal{E}$. □

The Canonical Model

The canonical model $\mathcal{M}^c = (W^c, E^c, R^c, \mathcal{E}^c, V^c)$

- $E^c = e \mid \varphi \mid \varphi \cdot \varphi \mid \varphi + \varphi$.
- $W^c = \{\langle \Gamma, F \rangle \mid \langle \Gamma, F \rangle \in \text{MCS} \times \mathcal{P}(E^c \times \mathit{Form})$, F satisfies the conditions as follows:
  (1) If $\langle s, \varphi \to \psi \rangle, \langle t, \varphi \rangle \in F$, then $\langle s \cdot t, \varphi \rangle \in F$;
  (2) If $\langle t, \varphi \wedge \psi \rangle \in F$, then $\langle t, \varphi \rangle \in F$ and $\langle t, \psi \rangle \in F$;
  (3) If $\langle s, \varphi \rangle \in F$ and $\langle t, \psi \rangle \in F$, then $\langle s + t, \varphi \wedge \psi \rangle \in F$;
  (4) If $\varphi \in \Gamma$ and φ is the formula such as $\mathcal{K}\psi$ or $\neg\mathcal{K}\psi$, then $\langle e, \varphi \rangle \in F$.
  (5) $\mathcal{K}_{why}\varphi \in \Gamma$ iff there exists t such that $\langle t, \varphi \rangle \in F\}$.

  MCS is the set of all maximal consistent sets in KW, and *Form* is the set of all formulas of KW. Following an established tradition, we denote elements of W^c as $\langle \Gamma, F \rangle, \langle \Delta, G \rangle$, and so forth;

- $\langle \Gamma, F \rangle R^c \langle \Delta, G \rangle$ iff $(\forall t \in E^c, \langle t, \varphi \rangle \in F$ iff $\langle t, \varphi \rangle \in G$ for each formula φ such that $\mathcal{K}_{why}\varphi \in \Gamma$, and $\Gamma^{\#} \subseteq \Delta$, where $\Gamma^{\#} = \{\mathcal{K}_{why}\varphi \mid \mathcal{K}_{why}\varphi \in \Gamma\} \cup \{\varphi \mid \mathcal{K}\varphi \in \Gamma\})$
- $\mathcal{E}^c(t, \varphi) = \{\langle \Gamma, F \rangle \mid \langle t, \varphi \rangle \in F\}$
- $V^c(p) = \{\langle \Gamma, F \rangle \mid p \in \Gamma\}$

$\mathcal{E}^c$ is well-defined. ($\mathcal{E}^c$ satisfies conditions (1)-(4) in the definition of $\mathcal{E}^c$.)

W^c is well-defined. (For any MCS Γ, there exists an F such that F satisfies conditions (1)-(5) in the definition of W^c.)

R^c is an equivalence relation.

To prove $\mathcal{E}^c$ is well-defined:

(1) Suppose $\langle \Gamma, F\rangle \in \mathcal{E}^c(s, \varphi \to \psi) \cap \mathcal{E}^c(t, \varphi)$. By the definition of $\mathcal{E}^c$, we have $\langle s, \varphi \to \psi\rangle, \langle t, \varphi\rangle \in F$. By the condition(1) of F in the definition of W^c, we have $\langle s \cdot t, \psi\rangle \in F$. Hence it follows that $\langle \Gamma, F\rangle \in \mathcal{E}^c(s \cdot t, \psi)$.

(2) Suppose $\langle \Gamma, F\rangle \in \mathcal{E}^c(t, \varphi \wedge \psi)$. By the definition of $\mathcal{E}^c$, we have $\langle t, \varphi \wedge \psi\rangle \in F$. By the condition(2) of F, we have $\langle t, \varphi\rangle, \langle t, \psi\rangle \in F$. Hence it follows that $\langle \Gamma, F\rangle \in \mathcal{E}^c\langle t, \varphi\rangle$ and $\langle \Gamma, F\rangle \in \mathcal{E}^c(t, \psi)$ by the definition of $\mathcal{E}^c$. Therefore $\langle \Gamma, F\rangle \in \mathcal{E}^c(t, \varphi) \cap \mathcal{E}^c(t, \psi)$.

(3) Suppose $\langle \Gamma, F\rangle \in \mathcal{E}^c(s, \varphi) \cap \mathcal{E}^c(t, \psi)$. By the definition of $\mathcal{E}^c$, we have $\langle s, \varphi\rangle, \langle t, \psi\rangle \in F$. By the condition(3) of F, we have $\langle s + t, \varphi \wedge \psi\rangle \in F$. Hence $\langle \Gamma, F\rangle \in \mathcal{E}^c(s + t, \varphi \wedge \psi)$.

Proposition

If $\langle\Gamma, F\rangle R^c\langle\Delta, G\rangle$, *then* $\mathcal{K}_{why}\varphi \in \Gamma$ *iff* $\mathcal{K}_{why}\varphi \in \Delta$ *and* $\mathcal{K}\varphi \in \Gamma$ *iff* $\mathcal{K}\varphi \in \Delta$.

Proof.

- Suppose $\mathcal{K}_{why}\varphi \in \Gamma$. By the definition of R^c, we have $\mathcal{K}_{why}\varphi \in \Delta$.
- Suppose $\mathcal{K}_{why}\varphi \in \Delta$ and $\mathcal{K}_{why}\varphi \notin \Gamma$. By the property of MCS, we have $\neg\mathcal{K}_{why}\varphi \in \Gamma$. By $\vdash \neg\mathcal{K}_{why}\varphi \to \mathcal{K}\neg\mathcal{K}_{why}\varphi$ and properties of MCS, we have $\mathcal{K}\neg\mathcal{K}_{why}\varphi \in \Gamma$. By the definition of R^c, we have $\neg\mathcal{K}_{why}\varphi \in \Delta$. Contradiction.
- Suppose $\mathcal{K}\varphi \in \Gamma$. By the axiom $\mathcal{K}\varphi \to \mathcal{K}\mathcal{K}\varphi$ and the property of MCS, we have that $\mathcal{K}\mathcal{K}\varphi \in \Gamma$. By the definition of R^c, we have that $\mathcal{K}\varphi \in \Delta$.
- Suppose $\mathcal{K}\varphi \in \Delta$ and $\mathcal{K}\varphi \notin \Gamma$. By the property of MCS, we have that $\neg\mathcal{K}\varphi \in \Gamma$. By the proposition $\neg\mathcal{K}\varphi \to \mathcal{K}\neg\mathcal{K}\varphi$ and the property

Proposition

For any MCS Γ, there exists an F such that F satisfies conditions (1)-(5) in the definition of W^c. (thus W^c is not empty.)

Proof.

Suppose Γ is an arbitrary maximal consistent set. We try to construct F as follows:

- $F_0 = \{\langle \varphi, \varphi \rangle \mid \mathcal{K}_{why}\varphi \in \Gamma\} \cup \{\langle e, \varphi \rangle \mid \varphi \in \Gamma$ and φ is the formula such as $\mathcal{K}\psi$ or $\neg\mathcal{K}\psi\}$.
- $F_1 = F_0 \cup \{\langle s \cdot t, \psi \rangle \mid \langle s, \varphi \to \psi \rangle \in F_0, \langle t, \varphi \rangle \in F_0\} \cup \{\langle t, \varphi \rangle \mid \langle t, \varphi \wedge \psi \rangle \in F_0\} \cup \{\langle t, \psi \rangle \mid \langle t, \varphi \wedge \psi \rangle \in F_0\} \cup \{\langle s + t, \varphi \wedge \psi \rangle \mid \langle s, \varphi \rangle \in F_0, \langle t, \psi \rangle \in F_0\}$.
- $\cdots\cdots$

Set $\Sigma = \{F_n \mid n \in N\}$ and $F = \bigcup \Sigma$. Obviously, by the construction of $F_n (n \in N)$, we have that Σ is monotonous.

To prove F is the set that satisfies conditions(1)-(5):

- Suppose $\langle s, \varphi \to \psi \rangle, \langle t, \varphi \rangle \in F$. By the monotonicity of Σ, there exists $n \in N$ such that $\langle s, \varphi \to \psi \rangle, \langle t, \varphi \rangle \in F_n$. Thus we have that $\langle s \cdot t, \psi \rangle \in F_{n+1}$ by the construction of $F_n (n \in N)$. Hence F satisfies condition (1).
- Suppose $\langle t, \varphi \wedge \psi \rangle \in F$. By the monotonicity of Σ, there exists $n \in N$ such that $\langle t, \varphi \wedge \psi \rangle \in F_n$. Thus we have that $\langle t, \varphi \rangle, \langle t, \psi \rangle \in F_{n+1}$ by the construction of $F_n (n \in N)$. Hence F satisfies condition (2).
- Suppose $\langle s, \varphi \rangle, \langle t, \psi \rangle \in F$. By the monotonicity of Σ, there exists $n \in N$ such that $\langle s, \varphi \rangle, \langle t, \psi \rangle \in F_n$. Thus we have that $\langle s + t, \varphi \wedge \psi \rangle \in F_{n+1}$ by the construction of $F_n (n \in N)$. Hence F satisfies condition (3).
- The constructions of F and F_0 guarantee that F satisfies condition (4).

Suppose $\mathcal{K}_{why}\varphi \in \Gamma$. Then we have that there exists φ such that $\langle \varphi, \varphi \rangle \in F$ by the construction of F_0 and F. On the other direction, we should show that $\forall t \in E^c$ and $\varphi \in \Gamma$, if $\langle t, \varphi \rangle \in F$, then $\mathcal{K}_{why}\varphi \in \Gamma$. Since $\langle t, \varphi \rangle \in F$, there exists $n \in N$ such that $\langle t, \varphi \rangle \in F_n$. We just need to show that $\forall n \in N$, if $\langle t, \varphi \rangle \in F_n$, then $\mathcal{K}_{why}\varphi \in \Gamma$. Use induction on n.

- Suppose $\langle t, \varphi \rangle \in F_0$. If $\langle t, \varphi \rangle \in \{\langle \varphi, \varphi \rangle \mid \mathcal{K}_{why}\varphi \in \Gamma\}$, obviously, we have $\mathcal{K}_{why}\varphi \in \Gamma$. If $\langle t, \varphi \rangle \in \{\langle e, \varphi \rangle \mid \varphi \in \Gamma$ and φ is the formula such as $\mathcal{K}\psi$ or $\neg\mathcal{K}\psi\}$, then we have $\varphi \in \Gamma$. By the properties of MCS and axioms $\mathcal{K}\varphi \rightarrow \mathcal{K}_{why}\mathcal{K}\varphi$ and $\neg\mathcal{K}\varphi \rightarrow \mathcal{K}_{why}\neg\mathcal{K}\varphi$, we have that $\mathcal{K}_{why}\varphi \in \Gamma$.
- Induction Hypothesis: $\forall t \in E^c$, if $\langle t, \varphi \rangle \in F_k$, then $\mathcal{K}_{why}\varphi \in \Gamma$. To prove that if $\langle t, \varphi \rangle \in F_{k+1}$, then $\mathcal{K}_{why}\varphi \in \Gamma$. Suppose $\langle t, \varphi \rangle \in F_{k+1}$.

- If $\langle t, \varphi \rangle \in F_k$, by IH, it can be proven.
- If $\langle t, \varphi \rangle \in \{\langle s \cdot t, \psi \rangle \mid \langle s, \varphi \to \psi \rangle \in F_k, \langle t, \varphi \rangle \in F_k\}$, then there exists ψ such that $\langle s, \psi \to \varphi \rangle \in F_k, \langle t, \psi \rangle \in F_k$. By IH, we have that $\mathcal{K}_{why}(\psi \to \varphi) \in \Gamma$ and $\mathcal{K}_{why}\psi \in \Gamma$. It follows that $\mathcal{K}_{why}\varphi \in \Gamma$ by the axiom $\mathcal{K}_{why}(\varphi \to \psi) \to (\mathcal{K}_{why}\varphi \to \mathcal{K}_{why}\psi)$ and the properties of MCS.
- If $\langle t, \varphi \rangle \in \{\langle t, \varphi \rangle \mid \langle t, \varphi \wedge \psi \rangle \in F_k\}$, then there exists ψ such that $\langle t, \varphi \wedge \psi \rangle \in F_k$. By IH, we have that $\mathcal{K}_{why}(\varphi \wedge \psi) \in \Gamma$. It follows that $\mathcal{K}_{why}\varphi \in \Gamma$ by the axiom $\mathcal{K}_{why}(\varphi \wedge \psi) \leftrightarrow \mathcal{K}_{why}\varphi \wedge \mathcal{K}_{why}\psi$ and the properties of MCS.
- If $\langle t, \varphi \rangle \in \{\langle t, \psi \rangle \mid \langle t, \varphi \wedge \psi \rangle \in F_k\}$. This situation is same as above.
- If $\langle t, \varphi \rangle \in \{\langle s + t, \varphi \wedge \psi \rangle \mid \langle s, \varphi \rangle \in F_k, \langle t, \psi \rangle \in F_k\}$, then there exists r, s, ψ and χ such that $t = r + s$, $\varphi = \psi \wedge \chi, \langle r, \psi \rangle \in F_k$ and $\langle s, \chi \rangle \in F_k$. By IH, we have that $\mathcal{K}_{why}\psi \in \Gamma$ and $\mathcal{K}_{why}\chi \in \Gamma$. Hence we have $\mathcal{K}_{why}(\psi \wedge \chi) \in \Gamma$ (i.e. $\mathcal{K}_{why}\varphi \in \Gamma$) by the axiom

Proposition

R^c is an equivalence relation.

Proof.

$\langle\Gamma, F\rangle R^c \langle\Delta, G\rangle$ iff (1) $\forall t \in E^c, \langle t, \varphi\rangle \in F$ iff $\langle t, \varphi\rangle \in G$ for each formula φ such that $\mathcal{K}_{why}\varphi \in \Gamma$, (2) $\Gamma^{\#} \subseteq \Delta$, where $\Gamma^{\#} = \{\mathcal{K}_{why}\varphi \mid \mathcal{K}_{why}\varphi \in \Gamma\} \cup \{\varphi \mid \mathcal{K}\varphi \in \Gamma\}$)

- Reflexivity: $\vdash \mathcal{K}\varphi \to \varphi$
- Transitivity: $\Vdash \mathcal{K}\psi \to \mathcal{K}\mathcal{K}\psi$
- Symmetry:

□

Lemma (Existence Lemma)

For the logic KW and any state $\langle \Gamma, F \rangle \in W^c$. If $\widehat{\mathcal{K}}\varphi \in \Gamma$ then there is a state $\langle \Delta, G \rangle \in W^c$ such that $\langle \Gamma, F \rangle R^c \langle \Delta, G \rangle$ and $\varphi \in \Delta$.

- Let Δ^- be $\{\varphi\} \cup \{\phi \mid \mathcal{K}\phi \in \Gamma\} \cup \{\mathcal{K}_{why}\psi \mid \mathcal{K}_{why}\psi \in \Gamma\}$.
- Δ^- is consistent. Suppose not.
  $\vdash_{KW} \phi_1 \wedge \cdots \wedge \phi_m \wedge \mathcal{K}_{why}\psi_1 \wedge \cdots \wedge \mathcal{K}_{why}\psi_n \to \neg\varphi$.
- $\vdash_{KW} \mathcal{K}(\phi_1 \wedge \cdots \wedge \phi_m \wedge \mathcal{K}_{why}\psi_1 \wedge \cdots \wedge \mathcal{K}_{why}\psi_n) \to \mathcal{K}\neg\varphi$
- $\vdash_{KW} (\mathcal{K}\phi_1 \wedge \cdots \wedge \mathcal{K}\phi_m \wedge \mathcal{K}\mathcal{K}_{why}\psi_1 \wedge \cdots \wedge \mathcal{K}\mathcal{K}_{why}\psi_n) \to \mathcal{K}\neg\varphi$.

It follows that $\mathcal{K}\neg\varphi \in \Gamma$. Using Dual, it follows that $\neg\widehat{\mathcal{K}}\varphi \in \Gamma$. But this is impossible: Γ is an MCS containing $\widehat{\mathcal{K}}\varphi$. We conclude that Δ^- is consistent.

Let Δ be any MCS containing Δ^-, such extensions exist by the Lindenbaum's Lemma. Let $G = F$. Obviously, $\langle \Delta, G \rangle$ satisfies conditions (1)-(3). We just need to check that $\langle \Delta, G \rangle$ satisfies conditions (4) and (5). Suppose $\varphi \in \Delta$ and φ is the formula such as $\mathcal{K}\psi$ or $\neg\mathcal{K}\psi$. We have two cases as follows:

- $\varphi = \mathcal{K}\psi$: Suppose $\mathcal{K}\psi \notin \Gamma$. By the property of MCS, we have that $\neg\mathcal{K}\psi \in \Gamma$. By the axiom $\neg\mathcal{K}\psi \to \mathcal{K}\neg\mathcal{K}\psi$ and the property of MCS, we have that $\mathcal{K}\neg\mathcal{K}\psi \in \Gamma$. Hence it follows that $\neg\mathcal{K}\psi \in \Delta$ by the construction of Δ. Contradiction.
- $\varphi = \neg\mathcal{K}\psi$: Suppose $\neg\mathcal{K}\psi \notin \Gamma$. By the property of MCS, we have that $\mathcal{K}\psi \in \Gamma$. By the axiom $\mathcal{K}\psi \to \mathcal{K}\mathcal{K}\psi$ and the property of MCS, we have that $\mathcal{K}\mathcal{K}\psi \in \Gamma$. Hence it follows that $\mathcal{K}\psi \in \Delta$ by the construction of Δ. Contradiction.

Hence we have that $\varphi \in \Gamma$. Since $G = F$, there exists $t \in E^c$, $\langle t, \varphi \rangle \in G$. We conclude that G satisfies condition (4).

$$\begin{array}{ll} & \text{For arbitrary } \mathcal{K}_{why}\varphi \in \Delta \\ \iff & \mathcal{K}_{why}\varphi \in \Gamma \text{ (by the construction of } \Delta) \\ \iff & \text{There exists } t \text{ such that } \langle t, \varphi \rangle \in F \text{ } (\langle \Gamma, F \rangle \text{ satisfies condition (5))} \\ \iff & \text{There exists } t \text{ such that } \langle t, \varphi \rangle \in G \text{ (G=F)} \end{array}$$

By construction of Δ and G, we have that $\varphi \in \Delta$ and $\langle \Gamma, F \rangle R^c \langle \Delta, G \rangle$.

Lemma (Truth Lemma)

For all φ'*s*, $\langle\Gamma, F\rangle \Vdash \varphi$ *if and only if* $\varphi \in \Gamma$

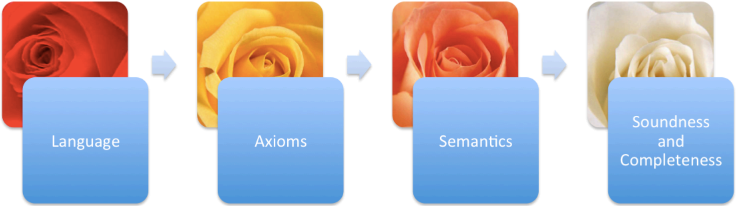
Proof.

This is established by standard induction on the complexity of φ. The atomic cases are covered by the definition of $'\Vdash'$. The Boolean induction steps are standard. Consider the case when φ is $\mathcal{K}_{why}\psi$ for some ψ.

$\Leftarrow$ If $\mathcal{K}_{why}\psi \in \Gamma$, for all $\langle\Delta, G\rangle$ such that $\langle\Gamma, F\rangle R^c \langle\Delta, G\rangle$, we have then $\mathcal{K}_{why}\psi \in \Delta$ by the definition of R^c. Since $\vdash \mathcal{K}_{why}\psi \to \psi$, we have $\psi \in \Delta$. By the Induction Hypothesis, $\langle\Delta, G\rangle \Vdash \psi$. By the condition(5) of F in the definition of W^c, we have that there exists $t \in E^c$ such that $\langle t, \psi\rangle \in F$. By the definition of R^c, we have that $\langle t, \psi\rangle \in G$. Hence $\langle\Delta, G\rangle \in \mathcal{E}^c(t, \psi)$ by the definition of $\mathcal{E}^c$. Therefore $\langle\Gamma, F\rangle \Vdash \mathcal{K}_{why}\psi$.

$\Rightarrow$ If $\mathcal{K}_{why}\psi \notin \Gamma$, then, by the condition(5) of W^c, we have

§5 Conditional Knowledge-Why Logic (abbr. CKW)



Language and Axioms

Language(single-agent): $\varphi = \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathcal{K}\varphi \mid \mathcal{K}_{why}(\varphi, \psi)$

- A1 Classical Propositional Axioms
- A2 $\mathcal{K}(\varphi \to \psi) \to (\mathcal{K}\varphi \to \mathcal{K}\psi)$
- A3 $\mathcal{K}_{why}(\alpha, \varphi \to \psi) \to (\mathcal{K}_{why}(\beta, \varphi) \to \mathcal{K}_{why}(\alpha \wedge \beta, \psi))$
- A4 $\mathcal{K}\varphi \to \varphi$
- A5 $\mathcal{K}_{why}(\alpha, \varphi) \to (\alpha \to \varphi)$
- A6 $\mathcal{K}\varphi \to \mathcal{K}\mathcal{K}\varphi$
- A7 $\neg\mathcal{K}\varphi \to \mathcal{K}\neg\mathcal{K}\varphi$
- A8 $\mathcal{K}\varphi \to \mathcal{K}_{why}(\top, \mathcal{K}\varphi)$
- A9 $\neg\mathcal{K}\varphi \to \mathcal{K}_{why}(\top, \neg\mathcal{K}\varphi)$
- A10 $\mathcal{K}_{why}(\alpha, \varphi) \to \mathcal{K}\mathcal{K}_{why}(\alpha, \varphi)$

Language and Axioms

A11 $\neg\mathcal{K}_{why}(\alpha,\varphi) \to \mathcal{K}\neg\mathcal{K}_{why}(\alpha,\varphi)$

A12 $\mathcal{K}_{why}(\alpha,\varphi\wedge\psi) \to \mathcal{K}_{why}(\alpha,\varphi)\wedge\mathcal{K}_{why}(\alpha,\psi)$

A13 $\mathcal{K}_{why}(\alpha,\varphi)\wedge\mathcal{K}_{why}(\beta,\psi) \to \mathcal{K}_{why}(\alpha\wedge\beta,\varphi\wedge\psi)$

A14 $\mathcal{K}_{why}(\bot,\varphi)$

A15 $\mathcal{K}(\alpha\to\beta) \to (\mathcal{K}_{why}(\beta,\varphi)\to\mathcal{K}_{why}(\alpha,\varphi))$

R1 Modus Ponens

R2 $\vdash\varphi \quad\Rightarrow\quad \vdash\mathcal{K}\varphi$

Semantics

item $w \Vdash \mathcal{K}_{why}(\alpha, \varphi)$ iff $(1) \exists t \in E, \forall v \in W, wRv,$ if $v \Vdash \alpha$, then $v \in \mathcal{E}(t, \varphi)$.

$(2) \forall v \in W, wRv,$ if $v \Vdash \alpha$, then $v \Vdash \varphi$.

Soundness

A15: $\mathcal{K}(\alpha \to \beta) \to (\mathcal{K}_{why}(\beta, \varphi) \to \mathcal{K}_{why}(\alpha, \varphi))$

Proof.

Suppose $w \Vdash_c \mathcal{K}(\varphi \to \psi)$, $w \Vdash_c \mathcal{K}_{why}(\psi, \chi)$ and $v \Vdash_c \varphi$ for all v such that wRv. Since $w \Vdash_c \mathcal{K}(\varphi \to \psi)$, we have that $v \Vdash_c \varphi \to \psi$ for all v such that wRv. Then we have that $v \Vdash_c \psi$ for all v such that wRv. Since $w \Vdash_c \mathcal{K}_{why}(\psi, \chi)$, by the definition of $\Vdash_c$, we therefore have that there exists t such that $v \in \mathcal{E}(t, \chi)$ for all v such that wRv. Hence we have $w \Vdash_c \mathcal{K}_{why}(\varphi, \chi)$. □

Completeness

$W^c = \{\langle \Gamma, F\rangle \mid \langle \Gamma, F\rangle \in$MCS$\times \mathcal{P}(E^c \times Form)$, F satisfies the conditions as follows:

(1) If $\langle \alpha, s, \varphi \to \psi\rangle, \langle \beta, t, \varphi\rangle \in F$, then $\langle \alpha \wedge \beta, s \cdot t, \varphi\rangle \in F$;

(2) If $\langle \alpha, t, \varphi \wedge \psi\rangle \in F$, then $\langle \alpha, t, \varphi\rangle \in F$ and $\langle \alpha, t, \psi\rangle \in F$;

(3) If $\langle \alpha, s, \varphi\rangle \in F$ and $\langle \alpha, t, \psi\rangle \in F$, then $\langle \alpha \wedge \beta, s + t, \varphi \wedge \psi\rangle \in F$;

(4) If α is an arbitrary formula, $\varphi \in \Gamma$ and φ is the formula such as $\mathcal{K}\psi$ or $\neg\mathcal{K}\psi$, then $\langle \alpha, e, \varphi\rangle \in F$.

(5) $\mathcal{K}_{why}(\alpha, \varphi) \in \Gamma$ iff there exists t such that $\langle \alpha, t, \varphi\rangle \in F\}$.

R^c $\langle \Gamma, F \rangle R^c \langle \Delta, G \rangle$ iff (1) If $\mathcal{K}_{why}(\alpha, \varphi) \in \Gamma$, then $\langle \alpha, t, \varphi \rangle \in F$ iff $\langle \alpha, t, \varphi \rangle \in G$. (2) $\Gamma^{\#} \subseteq \Delta$, where $\Gamma^{\#} = \{\mathcal{K}_{why}(\alpha, \varphi) \mid \mathcal{K}_{why}(\alpha, \varphi) \in \Gamma\} \cup \{\varphi \mid \mathcal{K}\varphi \in \Gamma\}$).

$\mathcal{E}$ $\mathcal{E}^c(t, \varphi) = \{\langle \Gamma, F \rangle \mid \langle \alpha, t, \varphi \rangle \in F, \alpha \in \Gamma\}$

ε^c is well-defined. (ε^c satisfies conditions (1)-(4) in the definition of ε^c.)

W^c is well-defined. (For any MCS Γ, there exists an F such that F satisfies conditions (1)-(5) in the definition of W^c.)

R^c is an equivalence relation.

§6 Future Work

[1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [15] [16] [14] [17] [18] [19] [20] [21] [22] [23] [24]

Reference I

[1] Artemov, S. (2008). The logic of justification. *The Review of Symbolic Logic*, 1(04):477–513.

[2] Balashov, Y. and Rosenberg, A. (2002). *Philosophy of science: contemporary readings*. Psychology Press.

[3] Bromberger, S. (1966). Questions. *The Journal of Philosophy*, pages 597–606.

[4] Colwell, G. (1996). Why-questions, determinism and circular reasoning. *Argumentation*, 10(1):1–24.

[5] Crawford, D. D. (1980). The cosmological argument, sufficient reason, and why-questions. *International Journal for Philosophy of Religion*, 11(2):111–122.

Reference II

[6] Cross, C. B. (1991). Explanation and the theory of questions. *Erkenntnis*, 34(2):237–260.

[7] Faye, J. (1999). Explanation explained. *Synthese*, 120(1):61–75.

[8] Fitting, M. (2005). The logic of proofs, semantically. *Annals of Pure and Applied Logic*, 132(1):1–25.

[9] Gabbay, D. M., Guenthner, F., and Guenthner, F. (2002). Handbook of philosophical logic.

[10] Gardenfors, P. (1980). A pragmatic approach to explanations. *Philosophy of science*, pages 404–423.

[11] Halonen, I. and Hintikka, J. (2005). Toward a theory of the process of explanation. *Synthese*, 143(1):5–61.

Reference III

[12] Hempel, C. G. and Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of science*, pages 135–175.

[13] Hintikka, J. and Halonen, I. (1995). Semantics and pragmatics for why-questions. *The Journal of Philosophy*, pages 636–657.

[14] Jaworski, W. (2009). The logic of how-questions. *Synthese*, 166(1):133–155.

[15] Kitcher, P. (1989). Explanatory unification and the causal structure of the world. *Scientific explanation*, 13:410–505.

[16] Koura, A. (1988). An approach to why-questions. *Synthese*, 74(2):191–206.

Reference IV

[17] Laland, K. N., Odling-Smee, J., Hoppitt, W., and Uller, T. (2013). More on how and why: cause and effect in biology revisited. *Biology & Philosophy*, 28(5):719–745.

[18] Lange, M. (2013). What makes a scientific explanation distinctively mathematical? *The British Journal for the Philosophy of Science*, 64(3):485–511.

[19] Markwick, P. (1999). Interrogatives and contrasts in explanation theory. *Philosophical studies*, 96(2):183–204.

[20] Nogina, E. (2014). On logic of formal provability and explicit proofs. *arXiv preprint arXiv:1405.2559*.

Reference V

[21] Sandborg, D. (1998). Mathematical explanation and the theory of why-questions. *The British Journal for the Philosophy of Science*, 49(4):603–624.

[22] Sintonen, M. (1984). On the logic of why-questions. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pages 168–176. JSTOR.

[23] Székely, G. (2011). *On why-questions in physics*. Springer.

[24] Weber, E., Van Bouwel, J., and De Vreese, L. (2013). *Scientific explanation*. Springer.