Occam's Razor by Minimal Negation

Satoshi Tojo[1] and Ken Satoh[2]

[1] Japan Advanced Institute of Science and Technology
tojo@jaist.ac.jp
[2] National Institute of Informatics
ksatoh@nii.ac.jp

Abstract. Historically *condito sine qua non* (c.s.q.n.) has been used as a criteria to judge if a person is culpable or not, though this criteria is known to be inadequate for some cases. Thus far, we have developed a theory to decide whether a certain statement is surely caused by the given antecedents in minimal abduction. However, the theory has included several problems concerning the meanings of logical connectives such as implication and logical-or. In this study, we employ *minimal negation* and revise these issues. We strictly distinguish the logical implication from causality, and consider the plausible semantics for the causation by Kripke semantics.

1 Introduction

This study is motivated by *condito sine qua non* (c.s.q.n. hereafter), which means 'if not' in English, that has worked as a principle to judge causality.

> "A caused B" if "if A had not happened, B would not have happened."
> (J. Glaser, 1858)

C.s.q.n. works as follows. Suppose that the lethal dose of a certain medicine is 100mg, and two persons A and B put 60mg of the medicine to a cup, respectively; when person C drank it, C died. If A had not put 60mg, C would not have died; thus, A is culpable for the death of C. So far so good; however, c.s.q.n. would turn to be inadequate in the following example. This time, the lethal dose is 100mg again, but each of A and B put 120mg to a cup, and naturally, C would die. If A had not put 120mg to the cup, would C have not die? No. Thus, A has no causal relation to the C's death in terms of c.s.q.n though A should be culpable.

C.s.q.n is related to another principle called *Occam's razor*, which states that 'the more reasons are employed, the less plausible the result becomes.' Namely, this principle implies the economy of reasoning, i.e., we need to find the minimal reason.

In this study, we are to propose a mechanism to find the minimal reason. But, in order to do this, we need to solve the inadequacy of such logical connective as negation, disjunction, and implication, as follows.

- Logical disjunction does not reflect the meaning of ‘or’ of natural language. We strictly distinguish the ‘or’ of meta-language from the logical connective.
- Negation in natural language should be eased to tolerate relative conflict.
- In legal reasoning, we often need to include multiple kinds of implications. We also distinguish the causal relation from the logical implication.

In our previous work, we clarified the distinction of meta-language ‘or’ and ‘$\vee$’ [1]. In this paper, we tackle a framework to decide the minimal reason for the causal relation with *Graded Negation* [2], together with the problems of logical ‘or’ and multiple implications.

2 C.S.Q.N. by Minimal explanation

We have tried to formalize the framework of minimal explanation by *minimal abduction* [1]. In this section, we briefly summarize the work. Given

$$\begin{cases} B \text{ Background theory} \\ H \text{ Abducibles (a set of propositional formulae)} \\ O \text{ A propositional formula} \end{cases}$$

we say E ($\subseteq H$) is an *explanation* if and only if

- $B \cup E \models O$ and $B \cup E \not\models \bot$.
- E is minimal if for any $E' \subset E$, $B \cup E' \not\models O$.

Now, let us consider:

Example 1 *There is a poison, the lethal dose of which is 100 mg, though below the lethal dose it gives some medical effect. A put 60 mg of it into C's coffee without telling C. Later, B also put another 60 mg into the same coffee without knowing A had already put the same poison. C drank the coffee and died.*

In this case, A and B's actions caused C's death since without either of actions, C would not have died. This example can be formalized in the following way in abduction.

$$\begin{cases} A_{60}\colon A \text{ put 60mg}, \quad B_{60}\colon B \text{ put 60mg} \\ C_{died}\colon C \text{ died} \\ B = \{A_{60} \wedge B_{60} \supset C_{died}\} \\ H_1 = \{A_{60},\ B_{60}\} \\ O = C_{died} \end{cases}$$

When $H_1 = \{A_{60},\ B_{60}\}$, the minimal explanations of O becomes $A_{60} \wedge B_{60}$ and this is the minimal cause.

Example 2 *A put 120 mg into C's coffee to kill C. B put another 120 mg into the same coffee to kill C without knowing A put the same poison. C drank the coffee and died.*

In this case, we cannot say that either of A and B's actions caused C's death; even if there weren't A's deed, C would have died. The second example is formalized as follows.

$$\begin{cases} A_{120}\colon A \text{ put 120mg}, \quad B_{120}\colon B \text{ put 120mg} \\ B = \{A_{120} \supset C_{died},\ B_{120} \supset C_{died},\ A_{120} \wedge B_{120} \supset C_{died}\} \\ H_2 = \{A_{120}, B_{120}\} \\ O = C_{died} \end{cases}$$

When $H_1 = \{A_{120},\ B_{120}\}$, the minimal explanations of O becomes $\{\{A_{120}\},\ \{B_{120}\}\}$. That is, either A_{120} or B_{120} is the minimal cause. In other words, there are two minimal explanations.

Example 3 *The inspector only found that A or B put 120mg into C's coffee. No one knows which of the two the true culprit is.*

The second example is formalized as follows.

$$\begin{cases} B = \{A_{120} \supset C_{died},\ B_{120} \supset C_{died}\} \\ H_3 = \{A_{120} \vee B_{120}\} \\ O = C_{died} \end{cases}$$

When $H_2 = \{A_{120} \vee B_{120}\}$, the minimal explanations of O becomes $\{\{A_{120} \vee B_{120}\}\}$. That is, $A_{120} \vee B_{120}$ is the minimal cause.

In this work, we have clarified the problem of *disjunction of causes* and *disjunctive causes*. However, the method of minimal explanation needs to generate a candidate explanation H' out of H rather in an ad hoc way, and does not include the process of cutting the redundant reasons off with the *razor*. In this study, we introduce the *comparison* process to decide which is less redundant.

Hereafter, to avoid notational confusion, we renounce the symbols of abduction, in which '$\models$' has been a valid proposition and '$\supset$' has been a logical implication. Instead, we employ '$\rightarrow$' for the logical implication.

3 Why graded negation?

Let us consider a situation that is caused by multiple reasons. For simplicity, δ is the consequence from the three precedents α, β, and γ.

$$\alpha \wedge \beta \wedge \gamma \rightarrow \delta.$$

Suppose that there is no other implication rule which produces δ. Then, if we can deny one of the three reasons, we can expect that δ is not deduced.

$$\neg\alpha \text{ negates } \alpha \wedge \beta \wedge \gamma.$$

However, in the similar way, we can deny multiple reasons.

$$\begin{array}{l} \neg\alpha \wedge \neg\beta \text{ negates } \alpha \wedge \beta \wedge \gamma \\ \neg\alpha \wedge \neg\beta \wedge \neg\gamma \text{ negates } \alpha \wedge \beta \wedge \gamma \end{array}$$

However, among $\neg\alpha$, $\neg\alpha \wedge \neg\beta$, and $\neg\alpha \wedge \neg\beta \wedge \neg\gamma$, the latter one more strongly negates $\alpha\wedge\beta\wedge\gamma$. This comparison can be applied to identify the *minimal reason*. In this discussion, either $\neg\alpha$, $\neg\beta$, or $\neg\gamma$ is the minimal reason to negate $\alpha\wedge\beta\wedge\gamma$.

Now, we briefly summarize *Graded negation* [2].

Definition 1 *Graded negation*

$$\Delta \vdash \neg_\alpha \beta \textit{ iff } \Delta \vdash \alpha \textit{ and } \alpha \wedge \beta \vdash \bot.$$

In this definition, α negates β in Δ; that is a weak negation, compared with the conventional negation. If $\Delta = \{\neg\beta, \gamma\}$ then $\Delta \vdash \neg\beta$ but $\Delta \nvdash \neg_\alpha\beta$.

Features of graded negation are summarized as follows.

- $\Delta \vdash \neg_\alpha\beta$ implies $\Delta \vdash \alpha \wedge \neg\beta$ because α does not appear from Δ.
- Because $\alpha \wedge \neg\alpha \vdash \bot$, $\neg_\alpha\neg\alpha \leftrightarrow \alpha$.
- Putting $\neg\alpha$ in the reverse way, we get $\neg_{\neg\alpha}\alpha \leftrightarrow \neg\alpha$.
- Given Δ, we cannot say $\Delta \vdash \neg_\alpha\beta \vee \neg_{\neg\alpha}\beta$ because neither α nor $\neg\alpha$ might appear from Δ.
- However, because β is compatible either with α or with $\neg\alpha$, $\vdash \neg_\alpha\beta \vee \neg\neg_\alpha\beta$.
- Also, $\Delta \vdash \neg\neg_\alpha\beta$ iff $\Delta \vdash \neg\alpha$ or $\{\alpha \wedge \beta\} \nvdash \bot$.

Definition 2 *Minimal negation*

$$\begin{array}{ll} \textit{α minimally negates δ} & \Delta \vdash \ominus_\alpha \delta, \\ \textit{iff} & \begin{cases} \Delta \vdash \neg_\alpha\delta, \textit{ and} \\ \textit{for any β such that } \Delta \vdash \neg_\beta\delta, \textit{ if } \vdash \alpha \to \beta, \textit{ then } \vdash \beta \to \alpha. \end{cases} \end{array}$$

For example, given $\Delta = \{\alpha, \beta, \gamma\}$, $\delta = \neg(\alpha \vee \beta)$,

$$\begin{cases} \Delta \vdash \neg_\alpha\delta, \\ \Delta \vdash \neg_{(\alpha\wedge\beta)}\delta, \\ \Delta \vdash \neg_{(\alpha\wedge\beta\wedge\gamma)}\delta. \end{cases}$$

Namely, each of α, $\alpha \wedge \beta$, and $\alpha \wedge \beta \wedge \gamma$ negates δ. However, the negation is tolerated as follows.

$$\begin{array}{ll} \neg_{(\alpha\wedge\beta)}\neg(\alpha\vee\beta) & : (\alpha\wedge\beta) \text{ strongly negates } \neg(\alpha\vee\beta) \\ \Downarrow & \\ \neg_\alpha\neg(\alpha\vee\beta) & : \alpha \text{ rather strongly negates } \neg(\alpha\vee\beta) \\ \Downarrow & \\ \neg_{(\alpha\vee\beta)}\neg(\alpha\vee\beta) & : (\alpha\vee\beta) \text{ weakly negates } \neg(\alpha\vee\beta) \end{array}$$

and $\alpha \vee \beta$ seems to negate δ most weakly. Now, we can assign truth values to the formulae with '$\ominus$' in accordance with Definition 2 in the following way.

$\ominus_{(\alpha\wedge\beta)}\delta$ F	$\neg_{(\alpha\wedge\beta)}\delta$ T	$(\alpha\wedge\beta)\to\alpha$ $\not\leftarrow$	$\neg_\alpha\delta$ T
$\ominus_\alpha\delta$ F	$\neg_\alpha\delta$ T	$\alpha\to(\alpha\vee\beta)$ $\not\leftarrow$	$\neg_{(\alpha\vee\beta)}\delta$ T
$\ominus_{(\alpha\vee\beta)}\delta$ T	$\neg_{(\alpha\vee\beta)}\delta$ T	$(\alpha\vee\beta)\to(\alpha\vee\beta\vee\gamma)$ $\not\leftarrow$	$\neg_{(\alpha\vee\beta\vee\gamma)}\delta$ F

Thus, we can conclude that

$$\neg_{(\alpha\vee\beta)}\neg(\alpha\vee\beta) \iff \ominus_{(\alpha\vee\beta)}\neg(\alpha\vee\beta),$$

that is, $\alpha \vee \beta$ minimally negates $\neg(\alpha \vee \beta)$.

In the following section, we develop this idea; we will contend that the minimal *reason* to negate some statement appears at the suffix of '$\ominus$'.

4 Causality, Minimality, and Culpability

4.1 Causality and Implication

In legal reasoning, there often happens a notational confusion originated from the mixture of causality and logical implication. For example, the rule of *weakening*

$$\frac{A \vdash C}{A \vee B \vdash C}$$

is generally admitted. But, the following inference

$$\begin{array}{ccc} A_{120} & \xrightarrow{caused} & C_{died} \\ & \Big\downarrow{\scriptstyle implies} & \\ A_{120} \vee B_{120} & \xrightarrow[caused?]{} & C_{died} \end{array}$$

is obviously inadequate. In any occasion, there is no reason that we weaken the cause of C_{died} from A_{120} to $A_{120} \vee B_{120}$. The transitivity between the causation and the logical implication, as well as the *deduction theorem*,

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash (A \to B)}$$

is meaningless if we stick to read '$\vdash$' as causality.

From here, we syntactically distinguish the logical implication '$\to$' from the causality. According to Lewis [6],

> 'α causes δ means that when α occurs δ is more likely than $\neg\delta$.'

We will restate this definition in Section 5 in terms of possible world semantics. As the first approximation, we state the definition in such a way that $\neg\delta$ deduces the inconsistency, as follows.

$$\Delta \vdash \alpha \quad and \quad \alpha \wedge \neg\delta \vdash \bot.$$

According to Definition 1, the above condition is exactly identical with $\Delta \vdash \neg_\alpha\neg\delta$.

Definition 3 *Causation by graded negation*
In a situation Δ, α causes δ iff $\Delta \vdash \neg_\alpha\neg\delta$.

Although $\alpha \vdash \delta$ implies $\alpha \wedge \neg\delta \vdash \bot$ and partially satisfies Definition 3, '$\vdash$' itself is no causal relation. We will discuss the adequacy of this definition precisely, later in Section 5.

Let us consider the example of overdosing, with the following set of general rules.

$$\Gamma_1 = \{A_{60} \wedge B_{60} \vdash C_{died},\ A_{120} \vdash C_{died},\ B_{120} \vdash C_{died}\}$$

As a preliminary step, we *circumscribe*[3] C_{died} from Γ_1.

$$C_{died} \leftrightarrow A_{120} \vee B_{120} \vee (A_{60} \wedge B_{60}).$$

In the above situation, when the amount of dose goes beyond 100mg, the tragedy would occur.

$$\begin{array}{l} \{A_{60}\} \nvdash C_{died}, \\ \{A_{60}, B_{60}\} \vdash C_{died}, \\ \{A_{120}\} \vdash C_{died}, \\ \{A_{120}, B_{120}\} \vdash C_{died}. \end{array}$$

Now, in the situation that both of A and B put 120mg of dose respectively, which is the minimal reason of the death of C? First, because $\{A_{120}, B_{120}\} \vdash A_{120} \wedge B_{120}$ and $A_{120} \wedge B_{120} \wedge C_{died} \vdash \bot$,

$$\{A_{120}, B_{120}\} \vdash \neg_{(A_{120} \wedge B_{120})} \neg C_{died}.$$

However, how can we tolerate the condition of C's death? Because in the similar reasoning process under $\{A_{120}, B_{120}\}$

$$\vdash \neg_{A_{120}} \neg C_{died},\ \ \vdash \neg_{B_{120}} \neg C_{died},\ \text{ and }\ \vdash \neg_{(A_{120} \vee B_{120})} \neg C_{died},$$

we can conclude that

$$\nvdash \ominus_{(A_{120} \wedge B_{120})} \neg C_{died},\ \ \nvdash \ominus_{A_{120}} \neg C_{died},\ \text{ and }\ \vdash \ominus_{(A_{120} \vee B_{120})} \neg C_{died}.$$

That is, $A_{120} \vee B_{120}$ is the minimal reason for C_{died}.

Thus, when we negate $\alpha \wedge \beta \wedge \gamma$, given a certain set of propositions Δ,

$$\begin{cases} \Delta \vdash \neg_{\neg\alpha}(\alpha \wedge \beta \wedge \gamma), \\ \Delta \vdash \neg_{(\neg\alpha \vee \neg\beta)}(\alpha \wedge \beta \wedge \gamma), \\ \Delta \vdash \ominus_{(\neg\alpha \vee \neg\beta \vee \neg\gamma)}(\alpha \wedge \beta \wedge \gamma). \end{cases}$$

However, we cannot tolerate the negation by arbitrary formula δ, as

$$\begin{cases} \Delta \vdash \neg_{\neg\alpha}(\alpha \wedge \beta \wedge \gamma), \\ \Delta \nvdash \neg_{(\neg\alpha \vee \delta)}(\alpha \wedge \beta \wedge \gamma), \end{cases}$$

because $\neg(\alpha \wedge \delta) \wedge (\alpha \wedge \beta \wedge \gamma) \nvdash \bot$.

[3] Given multiple clauses that defines φ, *circumscription* add such a new statement that '$\varphi \rightarrow \cdots$' [3]. This circumscription is not obligatory and is just for clarification. In this paper, we would like to explicitly show the disjunctive causes of C_{died} to be compared with the suffix of '$\neg$', to present how graded negation works.

4.2 Minimality and Culpability

The minimal reason is a logical consequence and can be the result of Occam's razor with which we could excise redundant reasons. However, this result does not always reflect if a person is culpable or not. For example, although

$$\{A_{120}, B_{120}\} \vdash \ominus_{(A_{120} \vee B_{120})} \neg C_{died},$$

this does not explain that the individual A is still accusable, regardless of B's deed of B_{120}. In order to judge if A (or B) is accusable or not, we need to provide the second step of judgement. For this purpose, we put an arbitrary proposition Q in a given situation Δ in the following place.

$$\Delta \vdash Q \wedge \neg C_{died}?$$

that is, 'are Q and $\neg C_{died}$ compatible in situation Δ?' This procedure is to put Q in the place of *cause* in Definition 3.

Now, we will revisit the examples of Section 2 and analyse them in the way of minimal negation.

Example 1 In this case, only when $A_{60} \wedge B_{60}$, C would be dead. Thus,

$$\{A_{60}, B_{60}\} \vdash \ominus_{(A_{60} \wedge B_{60})} \neg C_{died}.$$

How should the query $\vdash A_{60} \wedge \neg C_{died}$ be answered? Although $\{A_{60}, B_{60}\} \vdash A_{60}$, $\{A_{60}, B_{60}\} \vdash \neg_{A_{60}} \neg C_{died}$, i.e., as $A_{60} \wedge B_{60}$ is the minimal reason which cannot be tolerated to A_{60}. Thus, the answer to the query becomes *no*. Namely, the deed of A_{60} is a partial cause of C_{died}. However, A cannot be blamed unless (s)he knows B would put another 60mg. That is, we need to know the epistemic states of A and B, to judge if A is accusable or not.

Example 2 Because

$$\{A_{120}, B_{120}\} \vdash A_{120}, \ \vdash \ominus_{(A_{120} \vee B_{120})} \neg C_{died}, \text{ and } A_{120} \to A_{120} \vee B_{120},$$

we can conclude $\{A_{120}, B_{120}\} \vdash \neg_{A_{120}} \neg C_{died}$, that is, $\{A_{120}, B_{120}\} \not\vdash A_{120} \wedge \neg C_{died}$. A is accusable.

Example 3 In this case, the minimal reason becomes $A_{120} \vee B_{120}$ too, that is,

$$\{A_{120} \vee B_{120}\} \vdash \ominus_{(A_{120} \vee B_{120})} \neg C_{died}.$$

However, $\{A_{120} \vee B_{120}\} \not\vdash A_{120}$. That is, $\{A_{120} \vee B_{120}\} \not\vdash \neg_{A_{120}} \neg C_{died}$, i.e., A is not accusable yet.

5 Lewis' Causation and the Analysis of 'or'

When we cannot know either A or B put 120mg of medicine into the cup, we should write as

(A or B) put 120mg.

On the contrary, when we can obtain two possible solutions,

(A put 120mg) or (B put 120mg).

This opposition can be restated by the issue of scope of predicates. However, in the former case, we cannot provide 'or' inside a predicate as $put_{120}(A \text{ or } B)$.

One idea for us to represent such 'or' is to employ modal operators. The 'or' in knowledge, belief, and perception contexts is known to be decomposable [4]; e.g., "Ann knows that Betty or Chris won the race" does not imply "Ann knows that Betty won the race", nor "that Chris won the race." Employing a modal operator K_A meaning 'agent A knows that,' we can represent the distinction. In our case,

$$K_A(A_{120} \vee B_{120}) \not\Rightarrow K_A A_{120} \vee K_A B_{120}.$$

While the former claims that the cause is A_{120} or B_{120}, the latter says that either the cause is A_{120}, or the cause is B_{120}. With Kripke semantics (possible world semantics), $K_A\varphi$ holds in a certain world w if and only if φ holds in all the possible worlds accessible from w [5], given an accessibility relation R. Thus, the decomposability is explained as follows.

$$\begin{array}{rcl} w \models K_A(\varphi \vee \psi) & \Longleftrightarrow & \forall w'\ (wRw'),\ w' \models \varphi \text{ or } w' \models \psi \\ & & \not\Updownarrow \\ w \models K_A\varphi \vee K_A\psi & \Longleftrightarrow & \forall w'\ (wRw'),\ w' \models \varphi \text{ or } \forall w''\ (wRw''),\ w'' \models \psi \end{array}$$

The formalization by K_A may resolve the problem of epistemic state in **Example 1** as well as the decomposability of logical disjunction. However, to present a modal logic with full axiomatization and/or with full sequent system is far beyond the current scope. In this paper, we only utilize a part of possible world semantics concerning Lewis' causality, and discuss the future possibility later.

From a practical point of view, we introduce the *either-or* as follows. In a certain knowledge state w,

$$w \models (\varphi \bullet \psi \vdash \chi) \iff \text{either } w \models (\varphi \vdash \chi) \text{ or } w \models (\psi \vdash \chi). \tag{1}$$

Hereafter, we employ the possible world semantics of the above equation, to give the definition of causation [7, 8], that is,

> "$A \vdash B$ if and only if the world(s) in which $A \wedge B$ holds is closer to the real world than the world(s) in which $A \wedge \neg B$ holds, when possible worlds are partitioned by the plausibility."

Considering the conventional definition of logical disjunction ($\vee$)

$$w \models \varphi \vee \psi \iff w \models \varphi \text{ or } w \models \psi,$$

we develop the semantics of '$\bullet$' as follows.

$$w \models \varphi \bullet \psi \iff w \models \varphi \wedge \neg\psi \text{ or } w \models \neg\varphi \wedge \psi. \tag{2}$$

Now, given a usual triplet of Kripke frame $\langle W, R, \models\rangle$ where W is a set of possible worlds, R is the accessibility relation between worlds, and $\models$ is the valuation, we augment it to the quadruple $\langle W, R, \models, \leq\rangle$ where $\leq$ is the partition of possible worlds; i.e. $w_1 \leq w_2$ implies that w_1 is more plausible than w_2. Then, we give the semantics of '$\vdash$' according to Lewis' definition.

Definition 4 *Causation by plausibility* $\quad w \models (\alpha \vdash \beta)$ *if and only if*

$$\forall w' \ (wRw') \ w' \models \alpha \wedge \beta \ \textit{and} \ \forall w'' \ (wRw'') \ w'' \models \alpha \wedge \neg\beta, \ w' \leq w''.$$

Combining (2) and Definition 4, we obtain the semantics of

$$w \models (\varphi \bullet \psi \vdash \chi)$$

as follows.

$$\begin{array}{l}\forall w' \ (wRw') \ w' \models \varphi \wedge \neg\psi \text{ and } w' \models \chi \text{ or } w' \models \neg\varphi \wedge \psi \text{ and } w' \models \chi, \text{ and} \\ \forall w'' \ (wRw'') \ w'' \models \varphi \wedge \neg\psi \text{ and } w'' \models \neg\chi \text{ or } w'' \models \neg\varphi \wedge \psi \text{ and } w'' \models \neg\chi, \\ \ w' \leq w''.\end{array}$$

Namely,

$$\begin{array}{l}\forall w' \ (wRw') \ w' \models \varphi \wedge \neg\psi \wedge \chi \text{ or } w' \models \neg\varphi \wedge \psi \wedge \chi \text{ and} \\ \forall w'' \ (wRw'') \ w'' \models \varphi \wedge \neg\psi \wedge \neg\chi \text{ or } w'' \models \neg\varphi \wedge \psi \wedge \neg\chi, \\ \ w' \leq w''.\end{array} \tag{3}$$

At this stage, there have appeared four classes of possible worlds.

$$\begin{array}{l}w_1 \models \varphi \wedge \neg\psi \wedge \chi \\ w_2 \models \neg\varphi \wedge \psi \wedge \chi \\ w_3 \models \varphi \wedge \neg\psi \wedge \neg\chi \\ w_4 \models \neg\varphi \wedge \psi \wedge \neg\chi\end{array}$$

we can compare w_1 and w_3, and w_2 and w_4 in terms of the plausibility partition; however, we cannot do so between w_1 and w_4, and w_2 and w_3, because there are no common valuation among the three propositions. Therefore, (3) results in

$$\begin{array}{l}\forall w' \ (wRw') \ w' \models \varphi \wedge \chi \text{ and } \forall w'' \ (wRw'') \ w'' \models \varphi \wedge \neg\chi, \\ \quad w' \leq w'', \text{ or}\end{array}$$

$$\begin{array}{l}\forall w'' \ (wRw'') \ w'' \models \psi \wedge \chi \text{ and } \forall w'' \ (wRw'') \ w'' \models \psi \wedge \neg\chi, \\ \quad w' \leq w''.\end{array}$$

Getting back to Definition 4, we obtain

$$w \models (\varphi \vdash \chi) \text{ or } w \models (\psi \vdash \chi),$$

that is (1).

Below, we first revise the solution of **Example 3** using either-or, and then we introduce a more complicated example.

Example 3 *(revised) Either A or B put 120mg of the medicine into the coffee cup.*

The situation is formalized as

$$\{A_{120} \bullet B_{120}\} \vdash C_{died}.$$

According to (1),

$$\text{either } \{A_{120}\} \vdash C_{died} \text{ or } \{B_{120}\} \vdash C_{died},$$

and in each case, in some possible worlds w and w',

$$\begin{array}{l} \{A_{120}\} \vdash \ominus_{A_{120}} \neg C_{died} \text{ in } w \text{ or} \\ \{B_{120}\} \vdash \ominus_{B_{120}} \neg C_{died} \text{ in } w'. \end{array}$$

Example 4 *The inspector found that A put 60mg of the medicine to the C's cup, and in addition B put either 60mg or 120mg. C died.*

Now, we expand the general rules as follows.

$$\Gamma_2 = \Gamma_1 \cup \{A_{60} \wedge B_{120} \vdash C_{died}\},$$

when C_{died} is circumscribed as

$$C_{died} \leftrightarrow A_{120} \vee B_{120} \vee (A_{60} \wedge B_{60}) \vee (A_{60} \wedge B_{120}).$$

Then, the situation becomes

$$\{A_{60}, (B_{60} \bullet B_{120})\} \vdash C_{died}.$$

Because possible worlds are classified into the following two classes

$$w_1 \models A_{60} \wedge B_{60} \text{ and } w_2 \models A_{60} \wedge B_{120},$$

Thus,

$$A_{60} \wedge (B_{60} \bullet B_{120}) = (A_{60} \wedge B_{60}) \bullet (A_{60} \wedge B_{120}).$$

Namely, the either-or is decomposable in knowledge-base. This time, the minimal negation works as follows. Below, we omit the left-hand side of '$\vdash$'.

$$\begin{array}{l} \vdash \neg_{(A_{60} \wedge (B_{60} \vee B_{120}))} \neg C_{died}, \\ \vdash \neg_{(A_{60} \wedge B_{60})} \neg C_{died}, \\ \vdash \neg_{(A_{60} \wedge B_{120})} \neg C_{died}, \\ \vdash \neg_{B_{120}} \neg C_{died}. \end{array}$$

By the chains of logical strength

$$\begin{tikzcd} & A_{60} \wedge B_{120} \arrow[r] & B_{120} \\ A_{60} \wedge (B_{60} \bullet B_{120}) \arrow[ur] \arrow[dr] & & \\ & A_{60} \wedge B_{60} & \end{tikzcd}$$

only $(A_{60} \wedge B_{60})$ and B_{120} cannot be eased furthermore; thus, we can conclude

$$\vdash \ominus_{(A_{60} \wedge B_{60})} \neg C_{died}, \text{ and}$$
$$\vdash \ominus_{B_{120}} \neg C_{died}.$$

As this example, multiple minimal reasons may exist.

6 Conclusion

In this study, we have employed *minimal negation* and proposed a theory to find the minimal reason for a causation structure. The superability of this method is that we can compare two reasons directly and thus can put them in order as to which is stronger/weaker reasons.

The motivation of this study was how we could tune such logical connectives as negation, implication, and disjunction, to the meanings of practical reasoning. Employing the graded negation, we could relativize the notion of negation. We have defined the causality first by the minimal negation and have strictly distinguished the causal relation from the logical implication. We have discussed the adequacy of the definition in terms of possible world semantics (Kripke semantics), and by this, we have distinguished the semantics of 'or' in knowledge space from that of logical disjunction.

We have distinguished the minimality of reason from the culpability in Section 4.2. However, in order to judge if a person is accusable, we need to consider his/her knowledge state and/or intention. In Example 1, if B had known that A had put 60mg already B would be accusable; otherwise B would be innocent. As we have discussed in Section 5, we need to develop our theory to include such epistemic operators in future.

References

1. K. Satoh and S. Tojo. *Disjunction of Causes and Disjunctive Causes: a Solution to the Paradox of Condito Sine Qua None using Minimal Abduction*, JURIX'06, 2006.
2. D. Gabbay and A. Hunter. Negation and Contradiction in *What is negation?* (ed. D. Gabbay and H. Wansing), Kluwer Academic Publishers, 1999.
3. D. Gabbay, C. J. Hogger, and A. Robinson. (ed.) *Handbook of Logic in Artificial Intelligence*, Vol.3, Oxford Science Publications, 1995.
4. J. Barwise. *Situation in Logic*, CSLI lecture note, CSLI publications, 1989.
5. R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge*, The MIT Press, 1995.
6. D. R. Dowty. *Word Meaning and Montague Grammar*, Synthese Language Library, D. Reidel Publishing Company, 1979.
7. W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic states, in Harper et al. (ed.) *Causation in Decision, Belief Change, and Statistics vol.2*, pp. 105–134, Reidel, 1988.
8. W. Spohn. A general non-probabilistic theory of inductive reasoning, in Schachter et al. (ed.) *Uncertainty in Artificial Intelligence 4*, North-Holland, 1990.