# Rational Monotony in Input/Output Logic

**Xu Li** Liuwen Yu Leendert van der Torre

University of Luxembourg

DEON 2025, July 3, 2025
Vienna, Austria

# Contrary-to-duty paradoxes and SI

- Forrester's paradox:
  - Smith ought not to kill Jones. $(\top, \neg k)$
  - If Smith does kill Jones, then Smith ought to kill Jones gently. $(k, k \wedge g)$
  - Suppose that Smith kills Jones. $k$
- Strengthening of the Antecedent/Input (SI):

$$\frac{(a, x) \qquad b \vdash a}{(b, x)}$$

# Contrary-to-duty paradoxes and SI

- Forrester's paradox:
  - Smith ought not to kill Jones. $(\top, \neg k)$
  - If Smith does kill Jones, then Smith ought to kill Jones gently. $(k, k \wedge g)$
  - Suppose that Smith kills Jones. $k$
- Strengthening of the Antecedent/Input (SI):

$$\frac{(a, x) \qquad b \vdash a}{(b, x)}$$

$$\text{AND}\ \dfrac{\text{SI}\ \dfrac{(\top, \neg k)}{(k, \neg k)} \qquad (k, k \wedge g)}{(k, k \wedge \neg k \wedge g)}$$

UNIVERSITÉ DU LUXEMBOURG

# Introduction

- To deal with CTD paradoxes, SI must be weakened
- Our paper develops I/O logics where SI is replaced by (a form of) Rational Monotony (RM):

$$\neg \bigcirc (\neg\psi/\varphi) \to (\bigcirc(\chi/\varphi) \to \bigcirc(\chi/\varphi \wedge \psi))$$

  - if $\psi$ is permitted in context $\varphi$, then whatever is obligatory in context $\varphi$ is also obligatory in context $\varphi \wedge \psi$
  - $\dfrac{\varphi \nvdash \neg\psi, \varphi \vdash \chi}{\varphi \wedge \psi \vdash \chi}$ (Lehmann et al., 1992)
- Compared with constrained I/O logic, our approach provides better analysis of some CTD paradoxes

# I/O logic basics

- $\mathrm{PROP}$: finite non-empty set of atoms $\mathcal{L}$: propositional language on $\mathrm{PROP}$
- $A \vdash a$: $a$ is consequence of $A$ in PL $Cn(A)$: set of all consequences of $A$ in PL
- $a \dashv\vdash b$: $a$ is equivalent to $b$ in PL $Eq(a)$: set of all formulas equivalent to $a$
- $b \prec a$: $a \vdash b$ and $b \nvdash a$
- An output operation is a function $out : \wp(\mathcal{L} \times \mathcal{L}) \to \wp(\mathcal{L} \times \mathcal{L})$
  - A set $N \in \wp(\mathcal{L} \times \mathcal{L})$ is a **normative system**
    $(a, x) \in N$: given $a$, it ought to be the case that $x$
  - $out(N)$: set of (conditional) obligations that can be derived from $N$
  - $out(N, a) = \{x \mid (a, x) \in out(N)\}$
    collection of unconditional obligations in context $a$

# Properties of *out*

- Well-know properties of *out* (Makinson and Van Der Torre, 2000):

| | |
|---|---|
| REF | If $(a, x) \in N$, then $(a, x) \in out(N)$. |
| T | $(\top, \top) \in out(N)$. |
| SI | If $(a, x) \in out(N)$ and $b \vdash a$, then $(b, x) \in out(N)$. |
| WO | If $(a, x) \in out(N)$ and $x \vdash y$, then $(a, y) \in out(N)$. |
| AND | If $(a, x) \in out(N)$ and $(a, y) \in out(N)$, then $(a, x \wedge y) \in out(N)$. |
| OR | If $(a, x) \in out(N)$ and $(b, x) \in out(N)$, then $(a \vee b, x) \in out(N)$. |
| CT | If $(a, x) \in out(N)$ and $(a \wedge x, y) \in out(N)$, then $(a, y) \in out(N)$. |

# Four output operations in (Makinson and Van Der Torre, 2000)

$$\left.\begin{array}{l} out_1(N) \\ out_2(N) \\ out_3(N) \\ out_4(N) \end{array}\right\} \text{is the smallest set closed under} \left\{\begin{array}{l} \mathrm{REF, T, SI, WO, AND} \\ \mathrm{REF, T, SI, WO, AND, OR} \\ \mathrm{REF, T, SI, WO, AND, CT} \\ \mathrm{REF, T, SI, WO, AND, OR, CT} \end{array}\right.$$

- The representation results (semantics) for $out_1 - out_4$ are given in (Makinson and Van Der Torre, 2000)
- E.g., $out_1(N, a) = Cn(N(Cn(a)))$

# Dropping SI

## Definition

For each $1 \leq i \leq 4$, $out_i^-$ is the output operation obtained by substituting AT with T and IEQ with SI in the definition of $out_i$.

$$\begin{array}{rl} \text{IEQ} & \text{If } (a,x) \in out(N) \text{ and } a \dashv\vdash b, \text{ then } (b,x) \in out(N) \\ \text{AT} & (a,\top) \in out(N) \end{array}$$

Our paper gives the representation results for $out_1^- - out_3^-$.

# Example: Forrester's paradox

Let $N = \{(\top, \neg k), (k, k \wedge g)\}$. Then

- $out_1^-(N, a) = Cn(\neg k)$ if $a \dashv\vdash \top$;
- $out_1^-(N, a) = Cn(k \wedge g)$ if $a \dashv\vdash k$;
- $out_1^-(N, a) = Cn(\emptyset)$ if $a \not\dashv\vdash \top$ and $a \not\dashv\vdash k$.

# Example: Forrester's paradox

Let $N = \{(\top, \neg k), (k, k \wedge g)\}$. Then

- $out_1^-(N, a) = Cn(\neg k)$ if $a \dashv\vdash \top$;
- $out_1^-(N, a) = Cn(k \wedge g)$ if $a \dashv\vdash k$;
- $out_1^-(N, a) = Cn(\emptyset)$ if $a \not\dashv\vdash \top$ and $a \not\dashv\vdash k$.

Problem:

- Some meaningful conclusion not derived.
- Let $c$ be a proposition different to $k$ and $g$ (like "it is cloudy")
- Intuitively, given $c$, there is still the obligation not to kill
- However, $\neg k \notin out_1^-(N, c)$ as $c \not\dashv\vdash \top$

# Example: Forrester's paradox

Let $N = \{(\top, \neg k), (k, k \wedge g)\}$. Then

- $out_1^-(N, a) = Cn(\neg k)$ if $a \dashv\vdash \top$;
- $out_1^-(N, a) = Cn(k \wedge g)$ if $a \dashv\vdash k$;
- $out_1^-(N, a) = Cn(\emptyset)$ if $a \not\dashv\vdash \top$ and $a \not\dashv\vdash k$.

Problem:

- Some meaningful conclusion not derived.
- Let $c$ be a proposition different to $k$ and $g$ (like "it is cloudy")
- Intuitively, given $c$, there is still the obligation not to kill
- However, $\neg k \notin out_1^-(N, c)$ as $c \not\dashv\vdash \top$

" simply to drop SI is too heavy-handed. We need to know why SI is not always appropriate and, especially, when it remains justified" (Makinson and Torre, 2003)

UNIVERSITÉ DU LUXEMBOURG

# RM vs. wRM

- RM: If $(a, \neg b) \notin out(N)$ and $(a, x) \in out(N)$, then $(a \wedge b, x) \in out(N)$
- wRM: If $(a, \neg(a \wedge b)) \notin out(N)$ and $(a, x) \in out(N)$, then $(a \wedge b, x) \in out(N)$

# RM vs. wRM

- RM: If $(a, \neg b) \notin out(N)$ and $(a, x) \in out(N)$, then $(a \wedge b, x) \in out(N)$
- wRM: If $(a, \neg(a \wedge b)) \notin out(N)$ and $(a, x) \in out(N)$, then $(a \wedge b, x) \in out(N)$

## Proposition

- Let $out(N)$ be closed under WO. If $out(N)$ is closed under RM, then it is closed under wRM.
- Let $out(N)$ be closed under WO, AND and ID (given below). Then $out(N)$ is closed under RM iff it is closed under wRM.

$$\text{ID} \quad (a, a) \in out(N) \text{ for all formulas } a.$$

# RM vs. wRM

- RM: If $(a, \neg b) \notin out(N)$ and $(a, x) \in out(N)$, then $(a \wedge b, x) \in out(N)$
- wRM: If $(a, \neg(a \wedge b)) \notin out(N)$ and $(a, x) \in out(N)$, then $(a \wedge b, x) \in out(N)$
- To deal with CTD paradoxes, we will mainly focus on wRM
- Can we define, e.g., $out_1^{wr}(N)$ as the smallest set closed under $\{\mathrm{REF}, \mathrm{AT}, \mathrm{IEQ}, \mathrm{WO}, \mathrm{AND}, \mathrm{wRM}\}$?
- No. $out_1^{wr}(N)$ thus defined does not exist for certain $N$ (e.g., $N = \{(\top, c)\}$)
- Both RM and wRM are non-Horn rules.

UNIVERSITÉ DU LUXEMBOURG

# I/O logic as logical programs

- A logical program $\mathcal{P}$ is a set of rules of the form (where $a$ an atom and $l_i$ literals):
$$a \leftarrow l_1, \ldots, l_m$$
- A model for $\mathcal{P}$ is a valuation such that all rules in $\mathcal{P}$ are satisfied.
- If no negation appears in $\mathcal{P}$, then there exists an unique minimal model for $\mathcal{P}$. Otherwise, there might be multiple ones.
- Output operations as logical programs:

| $N$ | | $\mathcal{P}$ | |
|---|---|---|---|
| | | $\{(c, z) \leftarrow\} \cup$ | REF |
| $\{(c, z)\}$ | $\Rightarrow$ | $\{(b, y) \leftarrow (a, y) \mid b \vdash a\} \cup$ | SI |
| | | $\cdots \cup$ | $\ldots$ |
| | | $\{(a \wedge b, x) \leftarrow (a, x), \sim (a, \neg(a \wedge b))\}$ | wRM |

# Stable sets

## Definition (reduction)

Given a set $M \subseteq \mathcal{L} \times \mathcal{L}$, the *reduction* of wRM to $M$ is the following property $\text{wRM}|_M$:

$$\text{wRM}|_M \quad \begin{array}{l} \text{If } (a, \neg(a \wedge b)) \notin M \text{ and } (a, x) \in out(N), \\ \text{then } (a \wedge b, x) \in out(N). \end{array}$$

## Definition (stable set)

Let $N \subseteq \mathcal{L} \times \mathcal{L}$ and $\mathbb{P} \subseteq \{\text{REF}, \text{AT}, \text{IEQ}, \text{WO}, \text{AND}, \text{OR}, \text{CT}\}$. For all sets $M \subseteq \mathcal{L} \times \mathcal{L}$, let $out^M(N)$ be the smallest set closed under $\mathbb{P} \cup \{\text{wRM}|_M\}$. If $M = out^M(N)$, we say $M$ is a stable set of $N$ under $\mathbb{P} \cup \{\text{wRM}\}$.

UNIVERSITÉ DU LUXEMBOURG

# The representation result for $\mathbb{P}_1$

- We will focus on stable sets under the following four sets of properties $\mathbb{P}_1 - \mathbb{P}_4$:
  - $\mathbb{P}_1 = \{\mathsf{REF}, \mathsf{AT}, \mathsf{IEQ}, \mathsf{WO}, \mathsf{AND}\}$,
  - $\mathbb{P}_2 = \{\mathsf{REF}, \mathsf{AT}, \mathsf{IEQ}, \mathsf{WO}, \mathsf{AND}, \mathsf{OR}\}$,
  - $\mathbb{P}_3 = \{\mathsf{REF}, \mathsf{AT}, \mathsf{IEQ}, \mathsf{WO}, \mathsf{AND}, \mathsf{CT}\}$,
  - $\mathbb{P}_4 = \{\mathsf{REF}, \mathsf{AT}, \mathsf{IEQ}, \mathsf{WO}, \mathsf{AND}, \mathsf{OR}, \mathsf{CT}\}$.

UNIVERSITÉ DU LUXEMBOURG

# The representation result for $\mathbb{P}_1$

Let $N(A) = \{x \mid (a, x) \in N \text{ for some } a \in A\}$.

## Definition

Let $N \subseteq \mathcal{L} \times \mathcal{L}$. We define an output operation $out_1^{wr}(N)$ inductively as follows:

- If $a \dashv\vdash \top$, then $out_1^{wr}(N, a) = Cn(N(Eq(\top)))$;
- $out_1^{wr}(N, a) = Cn\left(N(Eq(a)) \cup \bigcup_{\{b \mid b \prec a \,\&\, \neg a \notin out_1^{wr}(N, b)\}} out_1^{wr}(N, b)\right)$.

## Theorem

*For all sets $N, M \subseteq \mathcal{L} \times \mathcal{L}$, $M$ is a stable set of $N$ under $\mathbb{P}_1 \cup \{\mathrm{wRM}\}$ iff $M = out_1^{wr}(N)$.*

UNIVERSITÉ DU LUXEMBOURG

# Example: Forrester's paradox

Let $N = \{(\top, \neg k), (k, k \wedge g)\}$. Then,

- $out_1^{wr}(N, a) = Cn(\neg k)$ if $a \nvdash k$;
- $out_1^{wr}(N, a) = Cn(k \wedge g)$ if $a \vdash k$ and $a \nvdash k \wedge \neg g$;
- $out_1^{wr}(N, a) = Cn(\emptyset)$ if $a \vdash k \wedge \neg g$.

# The cases $\mathbb{P}_2$–$\mathbb{P}_4$

## Definition

Let $N \subseteq \mathcal{L} \times \mathcal{L}$. For each $i \in \{2, 3, 4\}$, we define $out_i^{wr}(N)$ inductively as follows:

- if $a \dashv\vdash \top$, then $out_i^{wr}(N, a) = out_i^{-}(N, \top)$;
- $out_i^{wr}(N, a) = Cn\left(out_i^{-}(N, a) \cup \bigcup_{\{b \mid b \prec a \& \neg a \notin out_i^{wr}(N, b)\}} out_i^{wr}(N, b)\right)$.

In general, $out_i^{wr}(N)$ may not be a stable set of $N$ under the corresponding set of properties

# The cases $\mathbb{P}_2$–$\mathbb{P}_4$

## Proposition

For any set $N \subseteq \mathcal{L} \times \mathcal{L}$, the following hold:

- if $out_2^{wr}(N)$ is closed under OR, then $out_2^{wr}(N)$ is a stable set of $N$ under $\mathbb{P}_2 \cup \{\text{wRM}\}$.
- if $out_3^{wr}(N)$ is closed under CT, then $out_3^{wr}(N)$ is a stable set of $N$ under $\mathbb{P}_3 \cup \{\text{wRM}\}$.
- if $out_4^{wr}(N)$ is closed under OR and CT, then $out_4^{wr}(N)$ is a stable set of $N$ under $\mathbb{P}_4 \cup \{\text{wRM}\}$.

# Example: Chisholm's paradox

Let $N = \{(\top, g), (g, t), (\neg g, \neg t)\}$. Then:

- $out_3^{wr}(N, a) = Cn(g \wedge t)$ if $a \nvdash \neg g \vee \neg t$.
- $out_3^{wr}(N, a) = Cn(\emptyset)$ if $a \vdash \neg g \vee \neg t$ and $a \nvdash \neg g$.
- $out_3^{wr}(N, a) = Cn(\neg t)$ if $a \vdash \neg g$ and $a \nvdash \neg g \wedge t$.
- $out_3^{wr}(N, a) = Cn(\emptyset)$ if $a \vdash \neg g \wedge t$.

$out_3^{wr}(N)$ is a stable set of $N$ under $\mathbb{P}_3 \cup \{\text{wRM}\}$!

UNIVERSITÉ DU LUXEMBOURG

UNIVERSITÉ DU LUXEMBOURG

# Comparison with Constrained I/O Logic

- cIOL (Makinson and Torre, 2001) also intended to deal with CTD reasoning
- For Forrester's and Chisholm's paradoxes, our approach yields same result as cIOL
- But this does not hold in general:

## Example

Let $N = \{(\top, g), (\top, t), (\neg g, \neg t)\}$ and let the underlying unconstrained I/O logic be any of $out_1 - out_4$. We have:

$$out_c^{\cap}(N, \neg g, \neg g) = Cn(\emptyset)$$
$$out_c^{\cup}(N, \neg g, \neg g) = Cn(t) \cup Cn(\neg t)$$

In contrast, $out_1^{wr}(N, \neg g) = Cn(\neg t)$.

UNIVERSITÉ DU LUXEMBOURG

# Conclusion and Future Work

Summary:

- Two main approaches to deontic logic:
  - Preference-based: dyadic deontic logic
  - Rule-based: I/O logic
- This paper connects them by incorporating a key reasoning pattern into I/O logic: Rational Monotony
- "reduction" from stable semantics for logical programming/ASP

Future work:

- Stable sets in the cases of $\mathbb{P}_2 - \mathbb{P}_4$?
- Implementation in ASP?

UNIVERSITÉ DU LUXEMBOURG

# Reference


- Lehmann, Daniel and Menachem Magidor (1992). "What does a conditional knowledge base entail?" In: *Artificial Intelligence* 55.1, pp. 1–60.
- Makinson, David and Leendert van der Torre (Apr. 2001). "Constraints for Input/Output Logics". In: *Journal of Philosophical Logic* 30.2, pp. 155–185.
- — (2003). "What is Input/Output Logic?" In: *Foundations of the Formal Sciences II: Applications of Mathematical Logic in Philosophy and Linguistics, Papers of a Conference held in Bonn, November 10–13, 2000*. Ed. by Benedikt Löwe, Wolfgang Malzkom, and Thoralf Räsch. Dordrecht: Springer Netherlands, pp. 163–174.
- Makinson, David and Leendert Van Der Torre (2000). "Input/output logics". In: *Journal of philosophical logic* 29, pp. 383–408.