A Deontic Logic of Knowingly Complying

Carlos Areces, Valentin Cassano, Pablo Castro,
Raul Fervari and **Andrés R. Saravia**

CONICET, UNC and UNRC, Argentina
GTIIT, China

AAMAS, London – UK

31/V/2023

Motivation

Model a set of normative courses of actions

Motivation

Model a set of normative courses of actions and, given an initial and final conditions,

Motivation

Model a set of normative courses of actions and, given an initial and final conditions, whether the agents knowingly comply or not according to those courses.

DLKc: syntax & models

Model a set of normative courses of actions and, given an initial and final conditions, whether the agents knowingly comply or not according to those courses.

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \mathsf{N}(\psi, \varphi) \mid \mathsf{Kc}_i(\psi, \varphi) \text{ (i agent)}$$

DLKc: syntax & models

Model a set of normative courses of actions and, given an initial and final conditions, whether the agents knowingly comply or not according to those courses.

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \mathsf{N}(\psi, \varphi) \mid \mathsf{Kc}_i(\psi, \varphi) \text{ } (i \text{ agent})$$

$\mathfrak{M} = \langle \mathfrak{L}, \mathsf{U}, \mathsf{N} \rangle$ where:

DLKc: syntax & models

Model a set of normative courses of actions and, given an initial and final conditions, whether the agents knowingly comply or not according to those courses.

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \mathsf{N}(\psi, \varphi) \mid \mathsf{Kc}_i(\psi, \varphi) \text{ (i agent)}$$

$\mathfrak{M} = \langle \mathfrak{L}, \mathsf{U}, \mathsf{N} \rangle$ where:

- $\mathfrak{L} = \langle \mathsf{S}, \mathsf{R}, \mathsf{V} \rangle$ is an LTS;

DLKc: syntax & models

Model a set of normative courses of actions and, given an initial and final conditions, whether the agents knowingly comply or not according to those courses.

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \mathsf{N}(\psi, \varphi) \mid \mathsf{Kc}_i(\psi, \varphi) \text{ (i agent)}$$

$\mathfrak{M} = \langle \mathfrak{L}, \mathsf{U}, \mathsf{N} \rangle$ where:

- $\mathfrak{L} = \langle \mathsf{S}, \mathsf{R}, \mathsf{V} \rangle$ is an LTS;
- N a set of normative plans;

DLKc: syntax & models

Model a set of normative courses of actions and, given an initial and final conditions, whether the agents knowingly comply or not according to those courses.

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \mathsf{N}(\psi, \varphi) \mid \mathsf{Kc}_i(\psi, \varphi) \text{ (i agent)}$$

$\mathfrak{M} = \langle \mathfrak{L}, \mathsf{U}, \mathsf{N} \rangle$ where:

- $\mathfrak{L} = \langle \mathsf{S}, \mathsf{R}, \mathsf{V} \rangle$ is an LTS;
- N a set of normative plans;
- $\mathsf{U}(i)$ is an indistinguishability relation between plans for agent i.

DLKc: semantics

$\mathfrak{M} \Vdash \mathsf{N}(\psi, \varphi)$ iff exists a plan π such that

DLKc: semantics

$\mathfrak{M} \Vdash N(\psi, \varphi)$ iff exists a plan π such that

1. *is normative (i.e.* $\in$ N*),*

DLKc: semantics

$\mathfrak{M} \Vdash \mathsf{N}(\psi, \varphi)$ iff exists a plan π such that

1. *is normative (i.e. $\in$ N),*
2. *is fail-proof at all ψ-states*

DLKc: semantics

$\mathfrak{M} \Vdash N(\psi, \varphi)$ iff exists a plan π such that

1. *is normative (i.e. $\in$ N),*
2. *is fail-proof at all ψ-states and*
3. *always ends, from ψ-states, in φ-states.*

DLKc: semantics

$\mathfrak{M} \Vdash \mathsf{Kh}_i(\psi, \varphi)$ iff exists a set of plans $\Pi \in \mathsf{U}(i)$ such that every $\pi \in \Pi$

1. *is normative (i.e. $\in \mathsf{N}$),*
2. *is fail-proof at all ψ-states and*
3. *always ends, from ψ-states, in φ-states.*

DLKc: semantics

$\mathfrak{M} \Vdash \mathsf{Kh}_i(\psi, \varphi)$ iff exists a set of plans $\Pi \in U(i)$ such that every $\pi \in \Pi$

1. *is normative (i.e. $\in N$),*
2. *is fail-proof at all ψ-states and*
3. *always ends, from ψ-states, in φ-states.*

Note: $A\varphi = N(\neg\varphi, \bot)$ and $E\varphi = \neg A \neg\varphi$.

An example: Fire Emergency Evacuation Plan

EMERGENCY PROCEDURE

– FIRE – SMOKE – EXPLOSION	KEEP CALM PULL FIRE ALARM, FROM A SAFE LOCATION CALL 999 (FIRE BRIGADE)

Evacuate: close doors behind, use only stairs or ramps.

If unsafe to evacuate: shut door, block cracks, stay low near window.

An example: Fire Emergency Evacuation Plan

Emergency Procedure

– Fire – Smoke – Explosion	keep calm Pull Fire Alarm, from a safe location Call 999 (Fire Brigade)

Evacuate: close doors behind, use only stairs or ramps.

If unsafe to evacuate: shut door, block cracks, stay low near window.

An example: Fire Emergency Evacuation Plan

EMERGENCY PROCEDURE

– FIRE – SMOKE – EXPLOSION	KEEP CALM PULL FIRE ALARM, FROM A SAFE LOCATION CALL 999 (FIRE BRIGADE)

Evacuate: close doors behind, use only stairs or ramps.

If unsafe to evacuate: shut door, block cracks, stay low near window.

An example: Fire Emergency Evacuation Plan

Emergency Procedure

– Fire – Smoke – Explosion	keep calm Pull Fire Alarm, from a safe location Call 999 (Fire Brigade)

Evacuate: close doors behind, use only stairs or ramps (never elevator).

If unsafe to evacuate: shut door, block cracks, stay low near window.

An example: Fire Emergency Evacuation Plan

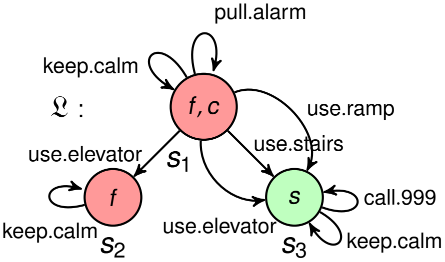
EMERGENCY PROCEDURE

– FIRE – SMOKE – EXPLOSION	KEEP CALM PULL FIRE ALARM, FROM A SAFE LOCATION CALL 999 (FIRE BRIGADE)

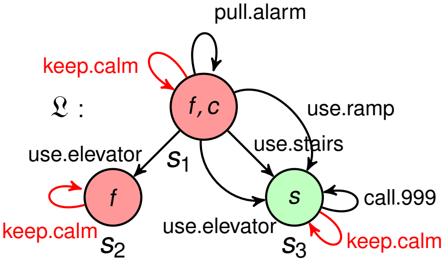
Evacuate: close doors behind, use only stairs or ramps.

If unsafe to evacuate: shut door, block cracks, stay low near window.

An example: An U-NLTS

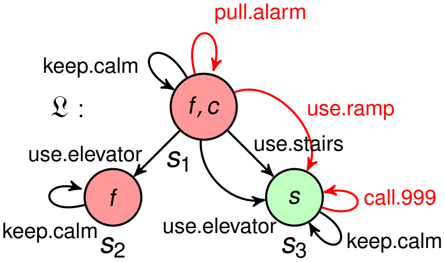


An example: An U-NLTS



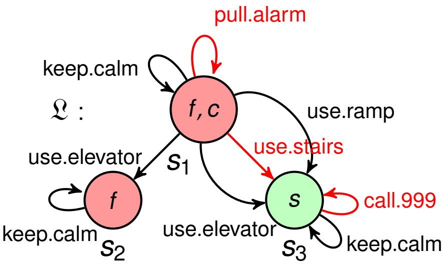
$$\pi_0 = \text{keep.calm}$$

An example: An U-NLTS



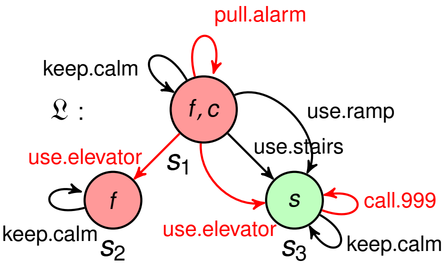
$\pi_0 =$ keep.calm
$\pi_r =$ pull.alarm; use.ramp; call.999

An example: An U-NLTS



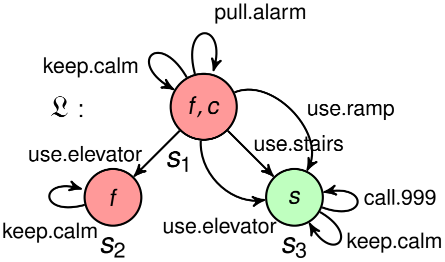
$\pi_0 =$ keep.calm
$\pi_r =$ pull.alarm; use.ramp; call.999
$\pi_s =$ pull.alarm; use.stairs; call.999

An example: An U-NLTS



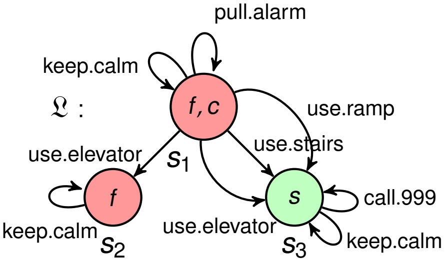
$\pi_0 =$ keep.calm
$\pi_r =$ pull.alarm; use.ramp; call.999
$\pi_s =$ pull.alarm; use.stairs; call.999
$\pi_e =$ pull.alarm; use.elevator; call.999

An example: An U-NLTS



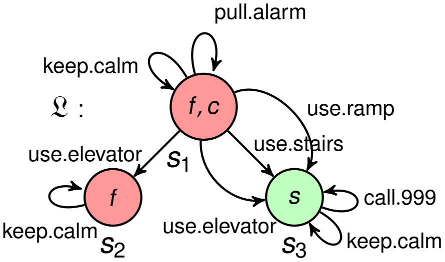
$\pi_0 =$ keep.calm
$\pi_r =$ pull.alarm; use.ramp; call.999
$\pi_s =$ pull.alarm; use.stairs; call.999
$\pi_e =$ pull.alarm; use.elevator; call.999
$N = \{\pi_0, \pi_r, \pi_s\}$

An example: An U-NLTS



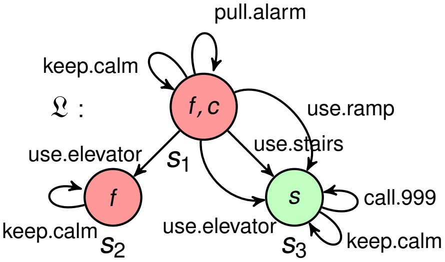
π_0 = keep.calm
π_r = pull.alarm; use.ramp; call.999
π_s = pull.alarm; use.stairs; call.999
π_e = pull.alarm; use.elevator; call.999
$N = \{\pi_0, \pi_r, \pi_s\}$

An example: An U-NLTS



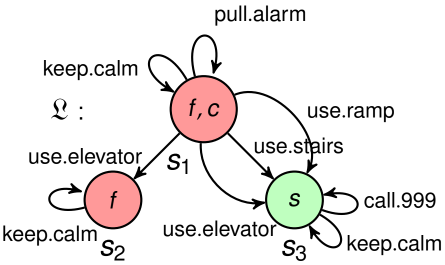
$\pi_0 = $ keep.calm
$\pi_r = $ pull.alarm; use.ramp; call.999
$\pi_s = $ pull.alarm; use.stairs; call.999
$\pi_e = $ pull.alarm; use.elevator; call.999
$\mathrm{N} = \{\pi_0, \pi_r, \pi_s\}$
$\mathrm{U}(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$

An example: An U-NLTS



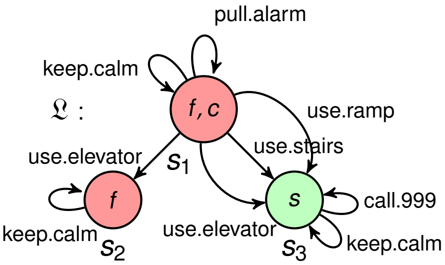
$$\pi_0 = \text{keep.calm}$$
$$\pi_r = \text{pull.alarm; use.ramp; call.999}$$
$$\pi_s = \text{pull.alarm; use.stairs; call.999}$$
$$\pi_e = \text{pull.alarm; use.elevator; call.999}$$
$$\text{N} = \{\pi_0, \pi_r, \pi_s\}$$
$$\text{U}(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$$
$$\text{U}(j) = \{\{\pi_e, \pi_r, \pi_s\}\}$$

An example: An U-NLTS



$\pi_0 = \text{keep.calm}$
$\pi_r = \text{pull.alarm; use.ramp; call.999}$
$\pi_s = \text{pull.alarm; use.stairs; call.999}$
$\pi_e = \text{pull.alarm; use.elevator; call.999}$
$\mathsf{N} = \{\pi_0, \pi_r, \pi_s\}$
$\mathsf{U}(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$
$\mathsf{U}(j) = \{\{\pi_e, \pi_r, \pi_s\}\}$
$\Vdash \mathsf{N}(f \wedge c, s)$

An example: An U-NLTS



$\pi_0 = \text{keep.calm}$

$\pi_r = \text{pull.alarm; use.ramp; call.999}$

$\pi_s = \text{pull.alarm; use.stairs; call.999}$

$\pi_e = \text{pull.alarm; use.elevator; call.999}$

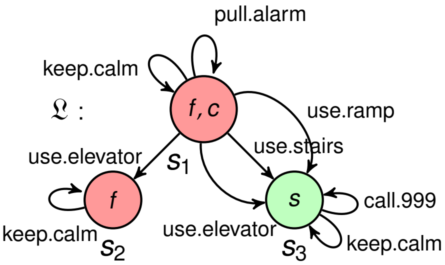
$\mathsf{N} = \{\pi_0, \pi_r, \pi_s\}$

$\mathsf{U}(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$

$\mathsf{U}(j) = \{\{\pi_e, \pi_r, \pi_s\}\}$

$\Vdash \mathsf{N}(f \wedge c, s)$

An example: An U-NLTS



$\pi_0 = \text{keep.calm}$

$\pi_r = \text{pull.alarm; use.ramp; call.999}$

$\pi_s = \text{pull.alarm; use.stairs; call.999}$

$\pi_e = \text{pull.alarm; use.elevator; call.999}$

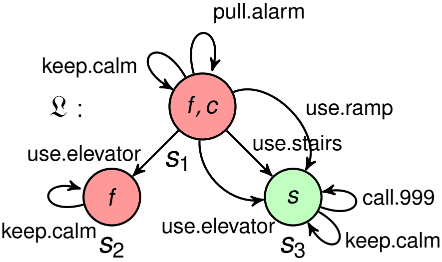
$\mathrm{N} = \{\pi_0, \pi_r, \pi_s\}$

$\mathrm{U}(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$

$\mathrm{U}(j) = \{\{\pi_e, \pi_r, \pi_s\}\}$

$\Vdash \mathrm{Kc}_i(f \wedge c, s)$

An example: An U-NLTS



π_0 = keep.calm

π_r = pull.alarm; use.ramp; call.999

π_s = pull.alarm; use.stairs; call.999

π_e = pull.alarm; use.elevator; call.999

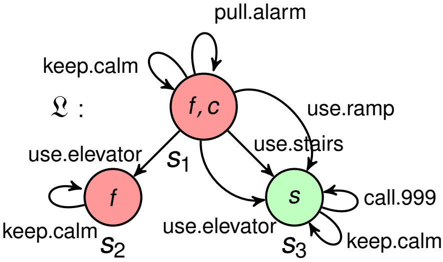
$N = \{\pi_0, \pi_r, \pi_s\}$

$U(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$

$U(j) = \{\{\pi_e, \pi_r, \pi_s\}\}$

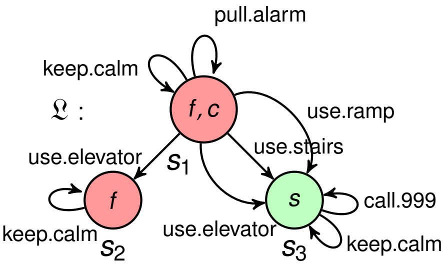
$\Vdash Kc_i(f \wedge c, s)$

An example: An U-NLTS



$\pi_0 = \text{keep.calm}$
$\pi_r = \text{pull.alarm; use.ramp; call.999}$
$\pi_s = \text{pull.alarm; use.stairs; call.999}$
$\pi_e = \text{pull.alarm; use.elevator; call.999}$
$\mathsf{N} = \{\pi_0, \pi_r, \pi_s\}$
$\mathsf{U}(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$
$\mathsf{U}(j) = \{\{\pi_e, \pi_r, \pi_s\}\}$
$\Vdash \neg \mathsf{Kc}_j(f \wedge c, s)$

An example: An U-NLTS



$\pi_0 = \text{keep.calm}$

$\pi_r = \text{pull.alarm; use.ramp; call.999}$

$\pi_s = \text{pull.alarm; use.stairs; call.999}$

$\pi_e = \text{pull.alarm; use.elevator; call.999}$

$\mathsf{N} = \{\pi_0, \pi_r, \pi_s\}$

$\mathsf{U}(i) = \{\{\pi_e\}, \{\pi_r, \pi_s\}\}$

$\mathsf{U}(j) = \{\{\pi_e, \pi_r, \pi_s\}\}$

$\Vdash \neg \mathsf{Kc}_j(f \wedge c, s)$

Axiom system for DLKc

Axioms:	
Taut	$\vdash \varphi$ for φ a propositional tautology
DistA	$\vdash \mathsf{A}(\psi \to \varphi) \to (\mathsf{A}\psi \to \mathsf{A}\varphi)$
TA	$\vdash \mathsf{A}\varphi \to \varphi$
4KcA	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\mathsf{Kc}_i(\psi, \varphi)$
5KcA	$\vdash \neg\mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi)$
4NA	$\vdash \mathsf{N}(\psi, \varphi) \to \mathsf{A}\,\mathsf{N}(\psi, \varphi)$
5NA	$\vdash \neg\,\mathsf{N}(\psi, \varphi) \to \mathsf{A}\neg\,\mathsf{N}(\psi, \varphi)$
KcN	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{N}(\psi, \varphi)$
DN	$\vdash \mathsf{N}(\varphi, \top)$
KcA	$\vdash (\mathsf{A}(\psi \to \chi) \wedge \mathsf{Kc}_i(\chi, \rho) \wedge \mathsf{A}(\rho \to \varphi)) \to \mathsf{Kc}_i(\psi, \varphi)$
NA	$\vdash (\mathsf{A}(\psi \to \chi) \wedge \mathsf{N}(\chi, \rho) \wedge \mathsf{A}(\rho \to \varphi)) \to \mathsf{N}(\psi, \varphi)$
Kc⊥	$\vdash \mathsf{Kc}_i(\bot, \bot)$
Rules:	
	$\dfrac{\vdash \psi \quad \vdash (\psi \to \varphi)}{\vdash \varphi}$ (MP) $\qquad \dfrac{\vdash \varphi}{\vdash \mathsf{A}\varphi}$ (Nec)

Axiom system for DLKc

Axioms:	
Taut	$\vdash \varphi$ for φ a propositional tautology
DistA	$\vdash \mathsf{A}(\psi \to \varphi) \to (\mathsf{A}\psi \to \mathsf{A}\varphi)$
TA	$\vdash \mathsf{A}\varphi \to \varphi$
4KcA	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\mathsf{Kc}_i(\psi, \varphi)$
5KcA	$\vdash \neg\mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi)$
4NA	$\vdash \mathsf{N}(\psi, \varphi) \to \mathsf{A}\,\mathsf{N}(\psi, \varphi)$
5NA	$\vdash \neg\,\mathsf{N}(\psi, \varphi) \to \mathsf{A}\neg\,\mathsf{N}(\psi, \varphi)$
KcN	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{N}(\psi, \varphi)$
DN	$\vdash \mathsf{N}(\varphi, \top)$
KcA	$\vdash (\mathsf{A}(\psi \to \chi) \land \mathsf{Kc}_i(\chi, \rho) \land \mathsf{A}(\rho \to \varphi)) \to \mathsf{Kc}_i(\psi, \varphi)$
NA	$\vdash (\mathsf{A}(\psi \to \chi) \land \mathsf{N}(\chi, \rho) \land \mathsf{A}(\rho \to \varphi)) \to \mathsf{N}(\psi, \varphi)$
Kc⊥	$\vdash \mathsf{Kc}_i(\bot, \bot)$

Rules:

$$\frac{\vdash \psi \quad \vdash (\psi \to \varphi)}{\vdash \varphi}\ (\mathrm{MP}) \qquad \frac{\vdash \varphi}{\vdash \mathsf{A}\varphi}\ (\mathrm{Nec})$$

Axiom system for DLKc

Axioms:	
Taut	$\vdash \varphi$ for φ a propositional tautology
DistA	$\vdash \mathsf{A}(\psi \to \varphi) \to (\mathsf{A}\psi \to \mathsf{A}\varphi)$
TA	$\vdash \mathsf{A}\varphi \to \varphi$
4KcA	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\mathsf{Kc}_i(\psi, \varphi)$
5KcA	$\vdash \neg\mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi)$
4NA	$\vdash \mathsf{N}(\psi, \varphi) \to \mathsf{A}\,\mathsf{N}(\psi, \varphi)$
5NA	$\vdash \neg\,\mathsf{N}(\psi, \varphi) \to \mathsf{A}\neg\,\mathsf{N}(\psi, \varphi)$
KcN	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{N}(\psi, \varphi)$
DN	$\vdash \mathsf{N}(\varphi, \top)$
KcA	$\vdash (\mathsf{A}(\psi \to \chi) \land \mathsf{Kc}_i(\chi, \rho) \land \mathsf{A}(\rho \to \varphi)) \to \mathsf{Kc}_i(\psi, \varphi)$
NA	$\vdash (\mathsf{A}(\psi \to \chi) \land \mathsf{N}(\chi, \rho) \land \mathsf{A}(\rho \to \varphi)) \to \mathsf{N}(\psi, \varphi)$
Kc⊥	$\vdash \mathsf{Kc}_i(\bot, \bot)$
Rules:	
	$\dfrac{\vdash \psi \quad \vdash (\psi \to \varphi)}{\vdash \varphi}$ (MP) $\qquad \dfrac{\vdash \varphi}{\vdash \mathsf{A}\varphi}$ (Nec)

Axiom system for DLKc

Axioms:	
Taut	$\vdash \varphi$ for φ a propositional tautology
DistA	$\vdash \mathsf{A}(\psi \to \varphi) \to (\mathsf{A}\psi \to \mathsf{A}\varphi)$
TA	$\vdash \mathsf{A}\varphi \to \varphi$
4KcA	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\mathsf{Kc}_i(\psi, \varphi)$
5KcA	$\vdash \neg\mathsf{Kc}_i(\psi, \varphi) \to \mathsf{A}\neg\mathsf{Kc}_i(\psi, \varphi)$
4NA	$\vdash \mathsf{N}(\psi, \varphi) \to \mathsf{A}\,\mathsf{N}(\psi, \varphi)$
5NA	$\vdash \neg\,\mathsf{N}(\psi, \varphi) \to \mathsf{A}\neg\,\mathsf{N}(\psi, \varphi)$
KcN	$\vdash \mathsf{Kc}_i(\psi, \varphi) \to \mathsf{N}(\psi, \varphi)$
DN	$\vdash \mathsf{N}(\varphi, \top)$
KcA	$\vdash (\mathsf{A}(\psi \to \chi) \land \mathsf{Kc}_i(\chi, \rho) \land \mathsf{A}(\rho \to \varphi)) \to \mathsf{Kc}_i(\psi, \varphi)$
NA	$\vdash (\mathsf{A}(\psi \to \chi) \land \mathsf{N}(\chi, \rho) \land \mathsf{A}(\rho \to \varphi)) \to \mathsf{N}(\psi, \varphi)$
Kc⊥	$\vdash \mathsf{Kc}_i(\bot, \bot)$
Rules:	
	$\dfrac{\vdash \psi \quad \vdash (\psi \to \varphi)}{\vdash \varphi}$ (MP) $\qquad \dfrac{\vdash \varphi}{\vdash \mathsf{A}\varphi}$ (Nec)

Properties:

- Sound and strongly complete axiom system.

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.
- A third modality (S(ψ, φ)):

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.
- A third modality (S(ψ, φ)):
  - General *abilities* of the agents.

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.
- A third modality (S(ψ, φ)):
  - General *abilities* of the agents.
  - *'what agents can do'* / *'what agents do according to norms'*.

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.
- A third modality (S(ψ, φ)):
  - General *abilities* of the agents.
  - *'what agents can do'* / *'what agents do according to norms'*.
  - Sound and strongly complete axiom system with the three.

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.
- A third modality (S(ψ, φ)):
  - General *abilities* of the agents.
  - *'what agents can do'* / *'what agents do according to norms'*.
  - Sound and strongly complete axiom system with the three.

Further work:

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.
- A third modality (S(ψ, φ)):
  - General *abilities* of the agents.
  - *'what agents can do'* / *'what agents do according to norms'*.
  - Sound and strongly complete axiom system with the three.

Further work:

- Study the complexity of the overall logic.

Properties:

- Sound and strongly complete axiom system.
- Model checking is in P.
- SAT problem is NP-complete.
- A third modality (S(ψ, φ)):
  - General *abilities* of the agents.
  - *'what agents can do'* / *'what agents do according to norms'*.
  - Sound and strongly complete axiom system with the three.

Further work:

- Study the complexity of the overall logic.
- Impose restrictions on the components of the model.