

# Ethical Reasoning for Autonomous Agents Under Uncertainty

**Michael Giancola**

Selmer Bringsjord

Naveen Sundar Govindarajulu

Carlos Varela



5th International Conference on  
Robot Ethics and Standards (ICRES 2020)

September 28, 2020



Sponsored by



# Motivation

# Motivation

- Autonomous agents are beginning to play significant roles in safety- and privacy-critical domains, such as driving and healthcare.

# Motivation

- Autonomous agents are beginning to play significant roles in safety- and privacy-critical domains, such as driving and healthcare.
- Humans are expected to be able to comprehend how to reason in these domains in order that their actions remain within any regulations/laws.

# Motivation

- Autonomous agents are beginning to play significant roles in safety- and privacy-critical domains, such as driving and healthcare.
- Humans are expected to be able to comprehend how to reason in these domains in order that their actions remain within any regulations/laws.
- We should expect the same (or more) from autonomous agents operating in these spaces.

# Motivation

- Autonomous agents are beginning to play significant roles in safety- and privacy-critical domains, such as driving and healthcare.
- Humans are expected to be able to comprehend how to reason in these domains in order that their actions remain within any regulations/laws.
- We should expect the same (or more) from autonomous agents operating in these spaces.
- Prior work is insufficient.

# Outline

# Outline

- Motivating prior work

# Outline

- Motivating prior work
- Deficiencies of the prior work

# Outline

- Motivating prior work
- Deficiencies of the prior work
- Necessary technical details

# Outline

- Motivating prior work
- Deficiencies of the prior work
- Necessary technical details
- Illuminating case study

# Outline

- Motivating prior work
- Deficiencies of the prior work
- Necessary technical details
- Illuminating case study
- Implementation

Dennis et al. (2016)  
Formal verification of ethical  
choices in autonomous systems

# Dennis et al. (2016)

## Formal verification of ethical choices in autonomous systems

- Presented a model-checking-based verification system for unmanned aircraft

# Dennis et al. (2016)

## Formal verification of ethical choices in autonomous systems

- Presented a model-checking-based verification system for unmanned aircraft
  - I. Determine the ethical violations to be considered

# Dennis et al. (2016)

## Formal verification of ethical choices in autonomous systems

- Presented a model-checking-based verification system for unmanned aircraft
  - I. Determine the ethical violations to be considered
  - 2. Rank them in order of increasing severity

# Dennis et al. (2016)

## Formal verification of ethical choices in autonomous systems

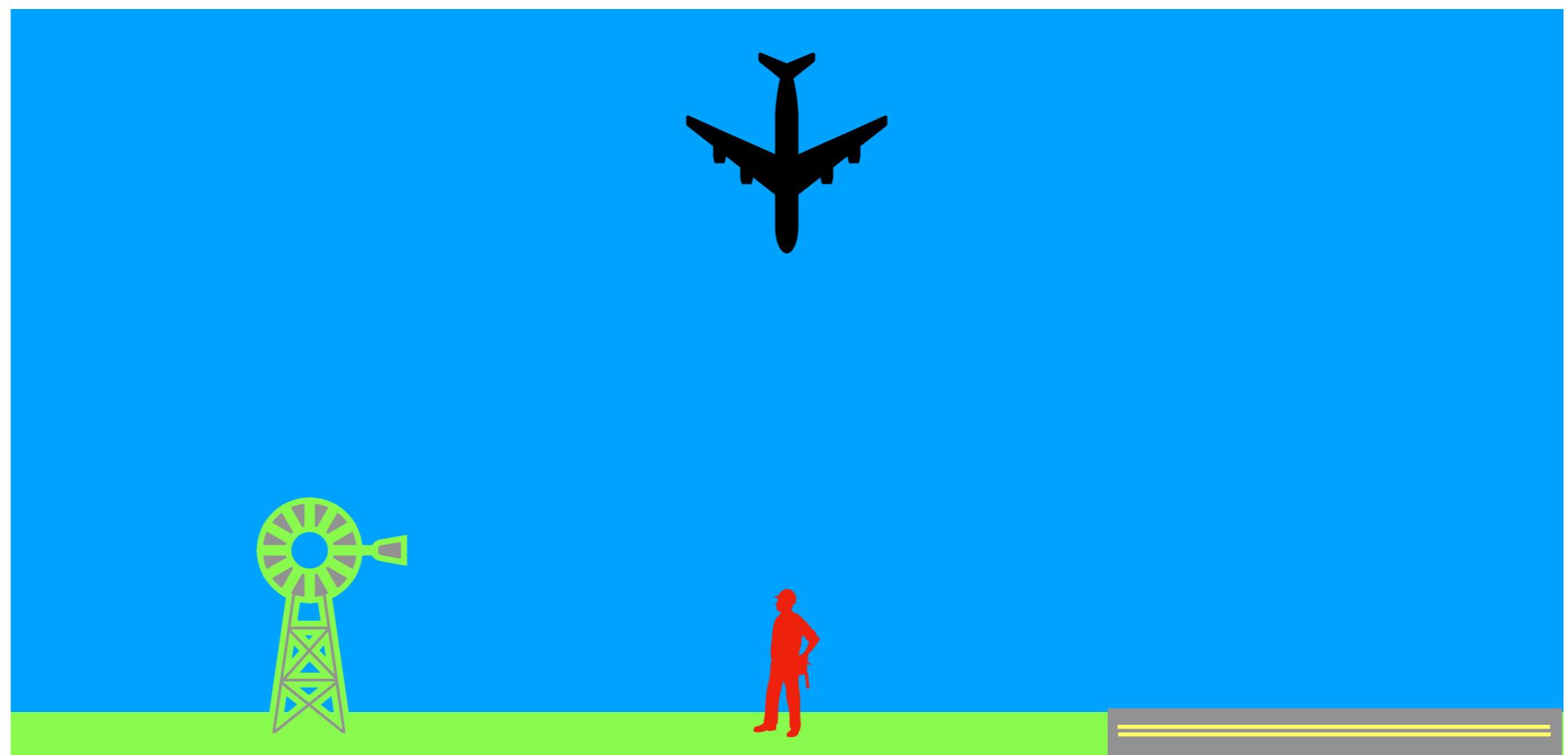
- Presented a model-checking-based verification system for unmanned aircraft
  - 1. Determine the ethical violations to be considered
  - 2. Rank them in order of increasing severity
  - 3. Choose the action which violates the least severe principle (or the fewest occurrences of said principle, as a tiebreaker)

# Dennis et al. (2016)

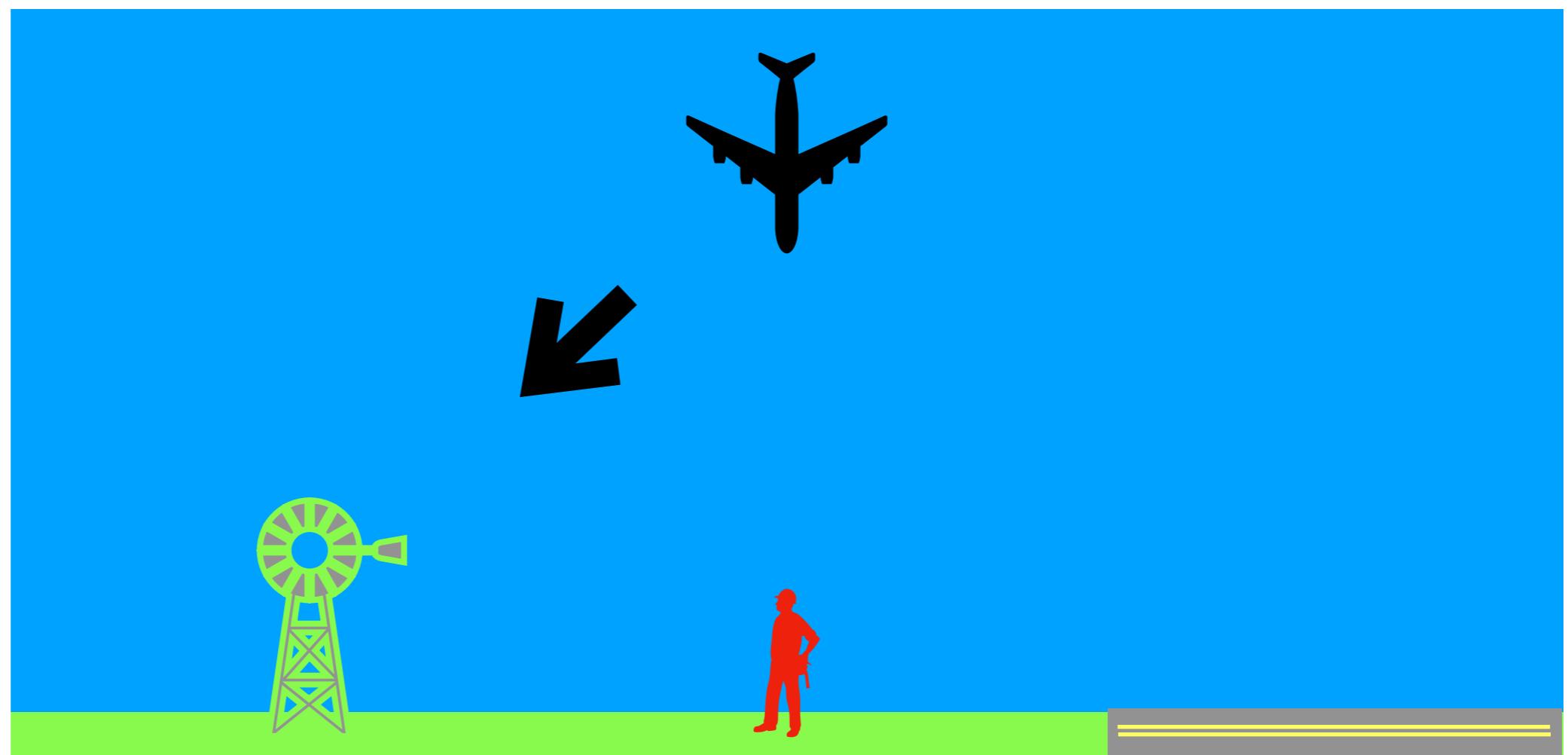
## Formal verification of ethical choices in autonomous systems

- Presented a model-checking-based verification system for unmanned aircraft
  - 1. Determine the ethical violations to be considered
  - 2. Rank them in order of increasing severity
  - 3. Choose the action which violates the least severe principle (or the fewest occurrences of said principle, as a tiebreaker)
- Described three potential scenarios and their solutions

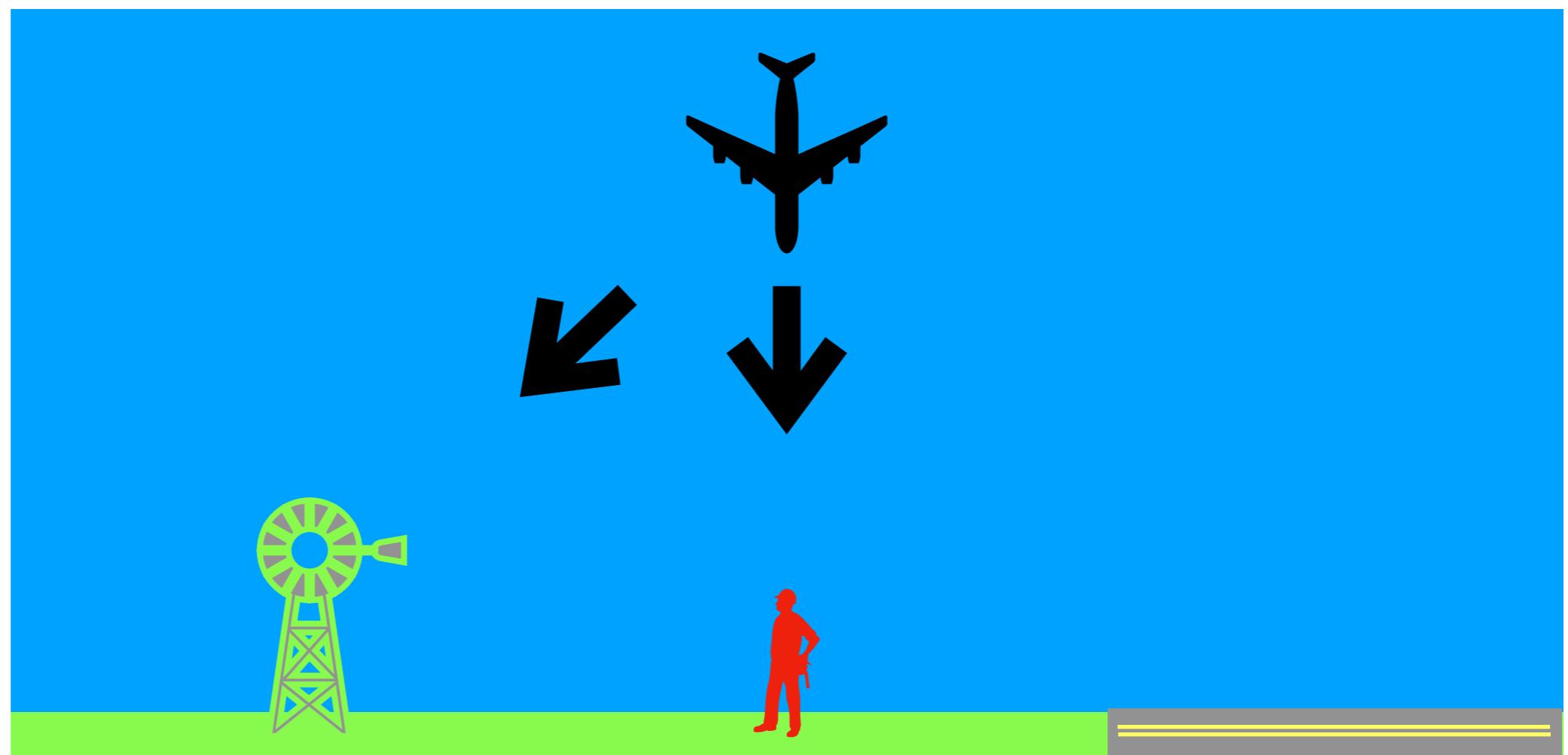
# Scenario III of Dennis et al. (2016): Fuel Low



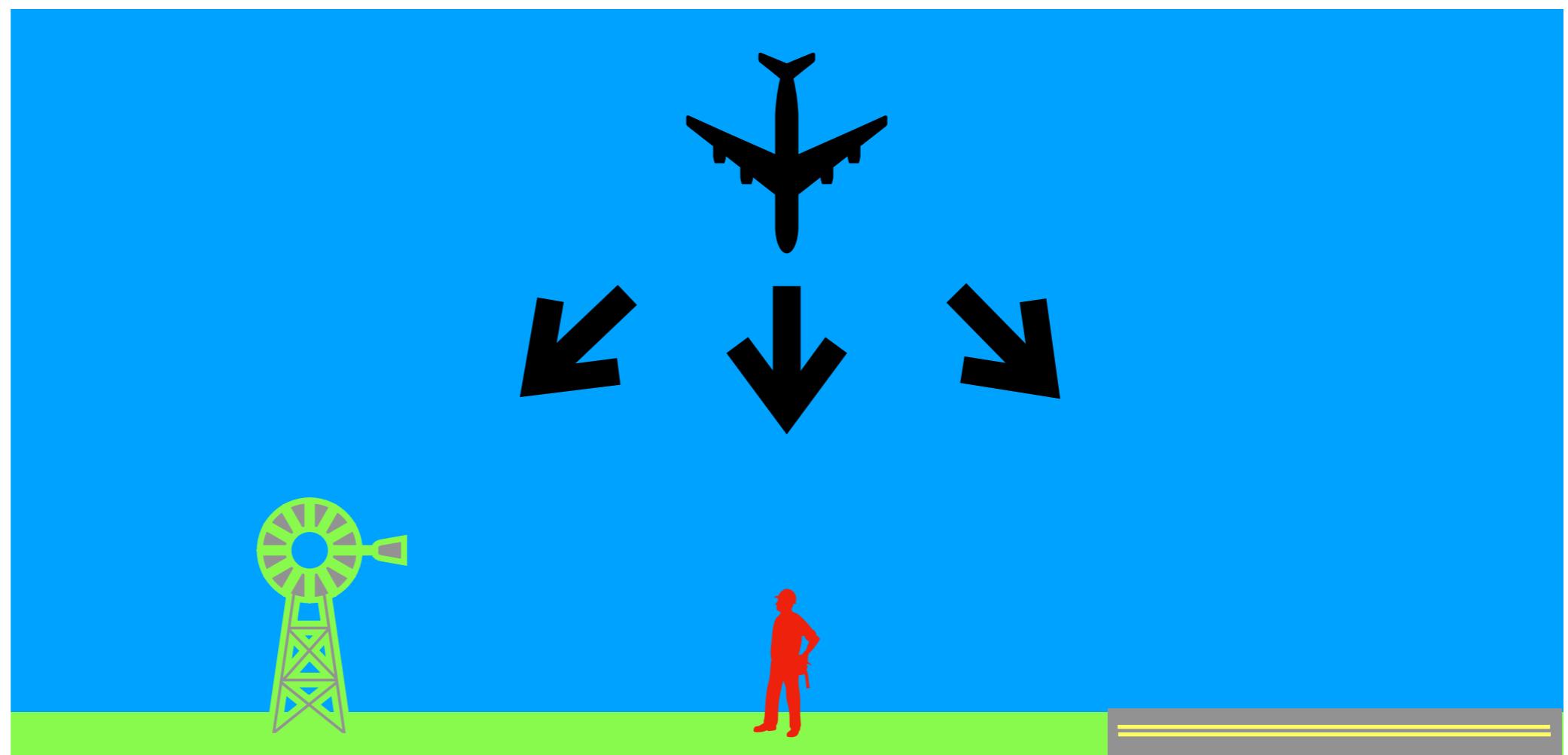
# Scenario III of Dennis et al. (2016): Fuel Low



# Scenario III of Dennis et al. (2016): Fuel Low



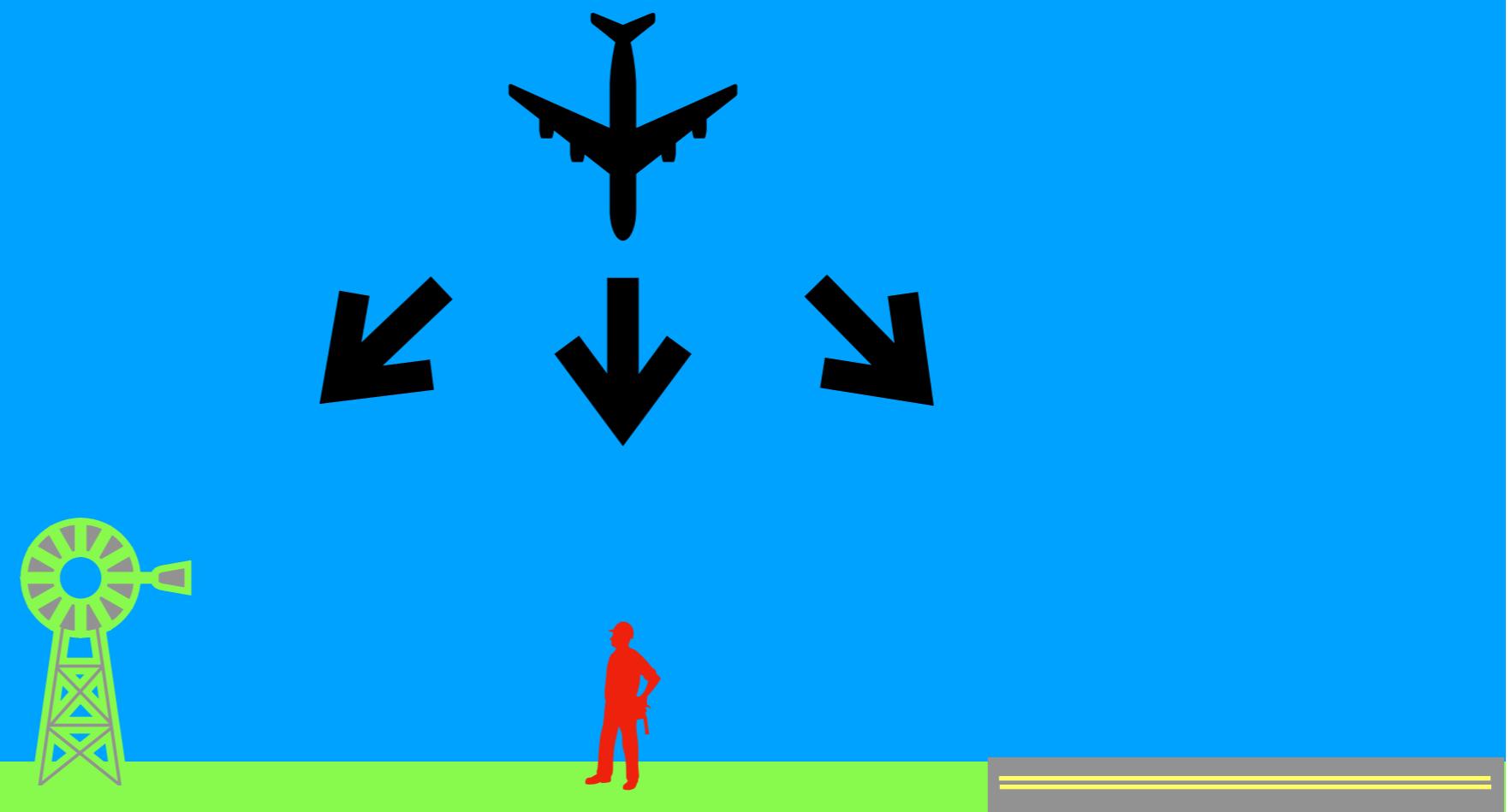
# Scenario III of Dennis et al. (2016): Fuel Low



# Scenario III of Dennis et al. (2016): Fuel Low

## Ethical Concerns:

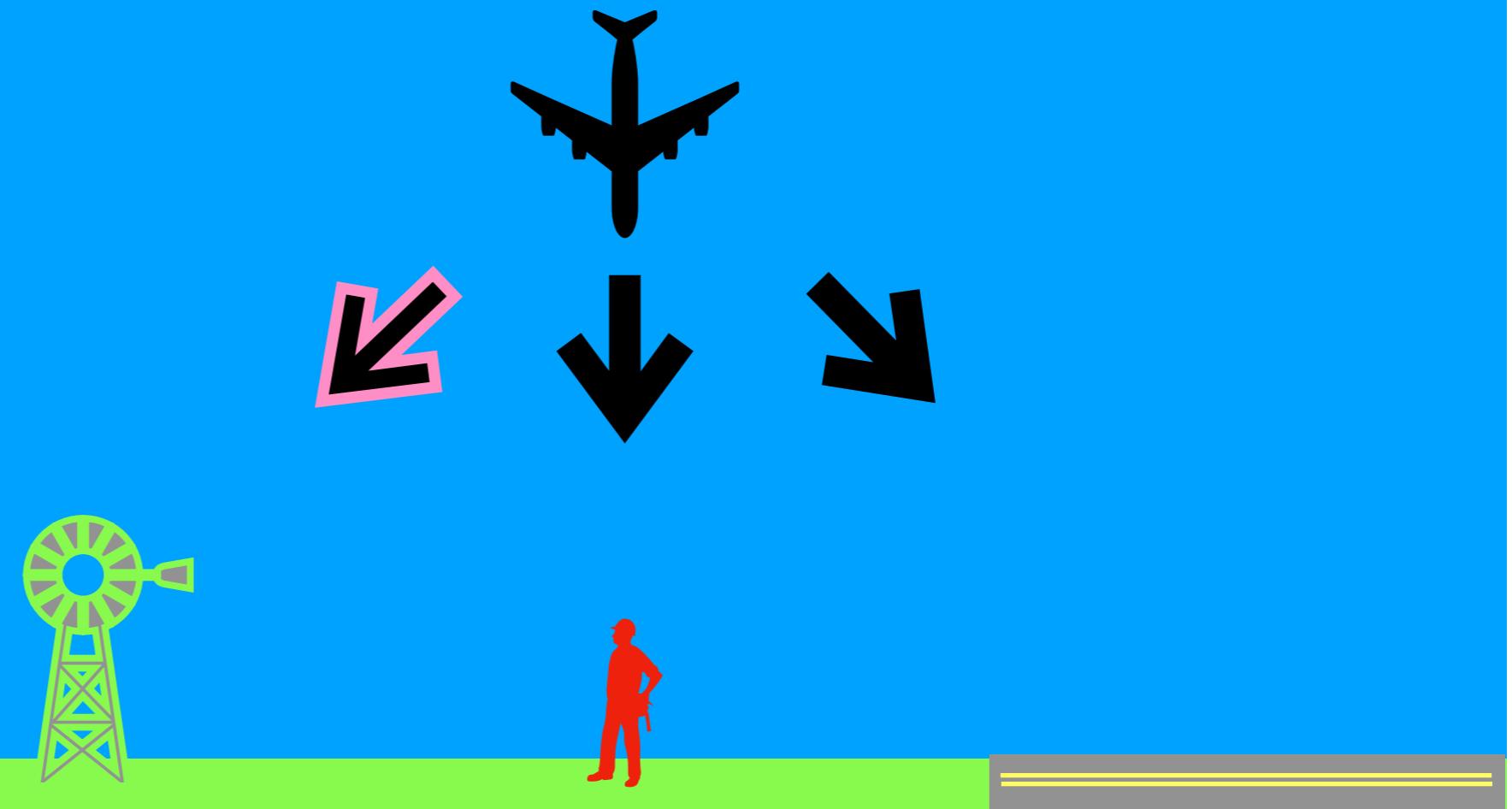
1. Do not damage own aircraft
2. Do not violate stay above 500 feet rule
3. Do not collide with objects on the ground
4. Do not cause damage to critical infrastructure
5. Do not collide with people



# Scenario III of Dennis et al. (2016): Fuel Low

## Ethical Concerns:

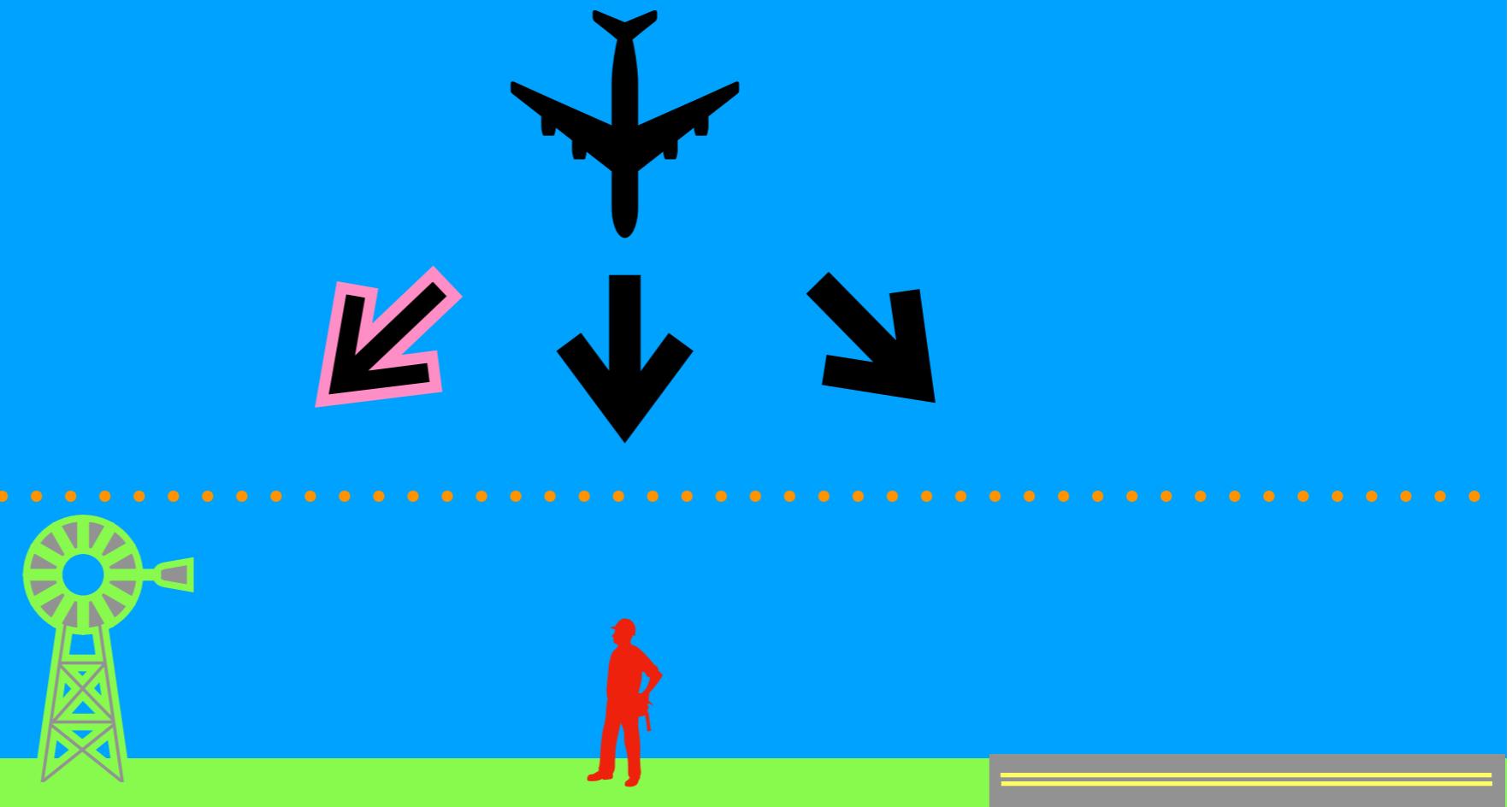
1. Do not damage own aircraft
2. Do not violate stay above 500 feet rule
3. Do not collide with objects on the ground
4. Do not cause damage to critical infrastructure
5. Do not collide with people



# Scenario III of Dennis et al. (2016): Fuel Low

## Ethical Concerns:

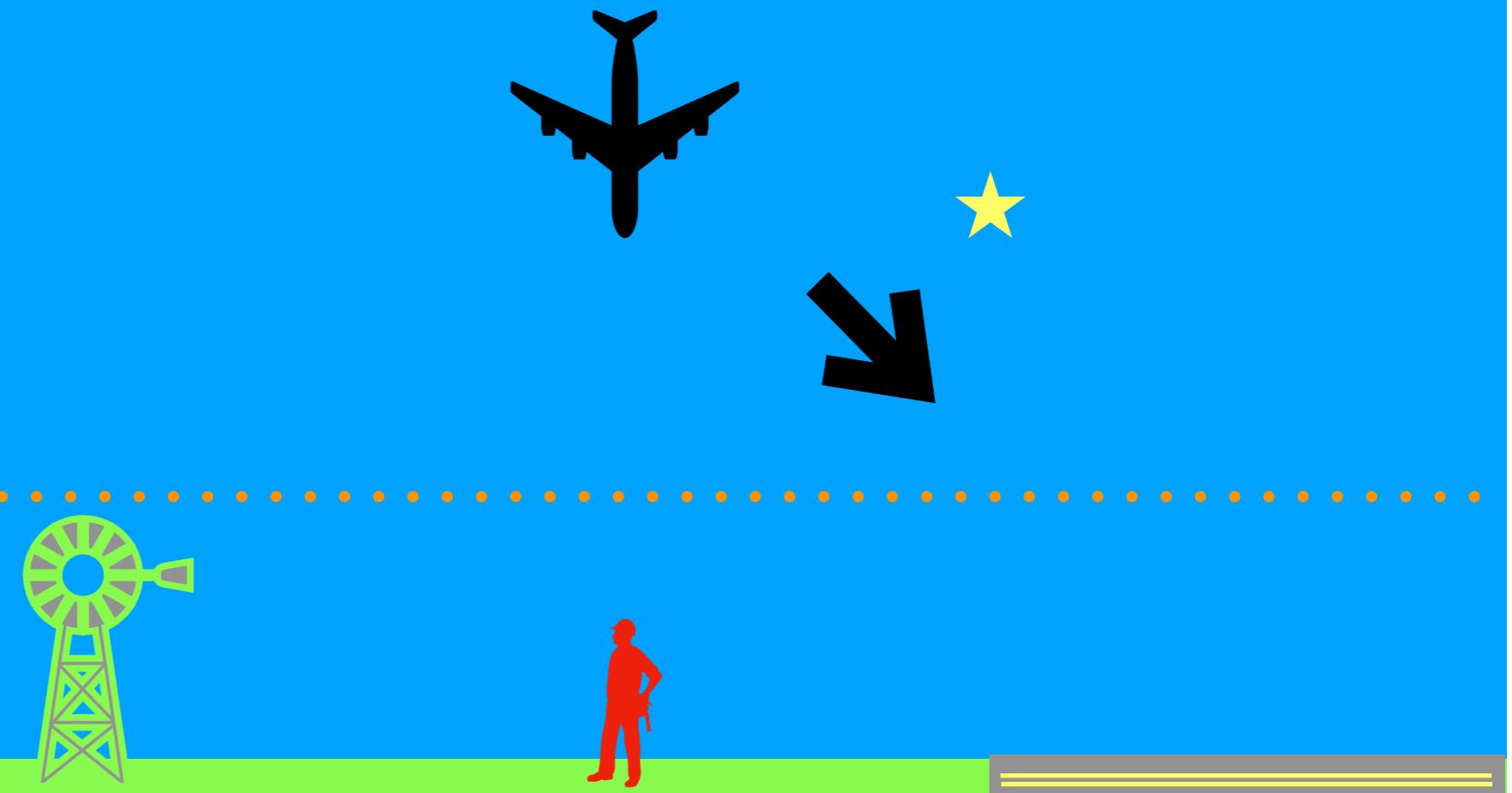
1. Do not damage own aircraft
2. Do not violate stay above 500 feet rule
3. Do not collide with objects on the ground
4. Do not cause damage to critical infrastructure
5. Do not collide with people



# Scenario III of Dennis et al. (2016): Fuel Low

## Ethical Concerns:

1. Do not damage own aircraft
2. Do not violate stay above 500 feet rule
3. Do not collide with objects on the ground
4. Do not cause damage to critical infrastructure
5. Do not collide with people



Flaws of Dennis  
et al. (2016)

&

Desiderata for  
Our Framework

Flaws of Dennis  
et al. (2016)

&

Desiderata for  
Our Framework

## I. Model checking is too slow

# Flaws of Dennis et al. (2016)

&

# Desiderata for Our Framework

I. Model checking is too slow

I. Verify ethical behavior on the  
order of seconds using a  
proof-theoretic approach

# Flaws of Dennis et al. (2016)

&

# Desiderata for Our Framework

- I. Model checking is too slow
- 2. BDI model is too expressive

- I. Verify ethical behavior on the order of seconds using a proof-theoretic approach

# Flaws of Dennis et al. (2016)

&

# Desiderata for Our Framework

- I. Model checking is too slow
- 2. BDI model is too inexpressive

- I. Verify ethical behavior on the order of seconds using a proof-theoretic approach
- 2. Highly expressive framework, capable of expressing nested belief & other modalities

# Flaws of Dennis et al. (2016)

&

# Desiderata for Our Framework

- I. Model checking is too slow
- 2. BDI model is too inexpressive
- 3. No use of uncertainty

- I. Verify ethical behavior on the order of seconds using a proof-theoretic approach
- 2. Highly expressive framework, capable of expressing nested belief & other modalities

# Flaws of Dennis et al. (2016)

&

# Desiderata for Our Framework

- 1. Model checking is too slow
- 2. BDI model is too inexpressive
- 3. No use of uncertainty

- 1. Verify ethical behavior on the order of seconds using a proof-theoretic approach
- 2. Highly expressive framework, capable of expressing nested belief & other modalities
- 3. Formalize uncertainty in order to reason about likelihood of ethical violations

# Technical Preliminaries

# Technical Preliminaries

- Cognitive Calculi

# Technical Preliminaries

- Cognitive Calculi
  - Multi-sorted FOL

# Technical Preliminaries

- Cognitive Calculi

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

**P**erceives, **K**nows, **S**ays, **C**ommon-knowledge  
**B**elieves, **D**esires, **I**ntends, **O**ught-to

# Technical Preliminaries

- Cognitive Calculi
- Multi-sorted FOL

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

# Technical Preliminaries

- Cognitive Calculi
  - Multi-sorted FOL
  - Inference Schemata

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

# Technical Preliminaries

## Inference Schemata

$$\begin{array}{c}
 \frac{}{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \rightarrow \mathbf{K}(a, t, \phi))} [I_1] \quad \frac{}{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \rightarrow \mathbf{B}(a, t, \phi))} [I_2] \\
 \frac{\mathbf{K}(a, t, \phi)}{\phi} [I_4] \\
 \frac{\mathbf{S}(s, h, t, \phi)}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} [I_{12}] \quad \frac{\mathbf{I}(a, t, \text{happens}(\text{action}(a^*, \alpha), t'))}{\mathbf{P}(a, t, \text{happens}(\text{action}(a^*, \alpha), t'))} [I_{13}] \\
 \frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} [I_{14}] \\
 \frac{\mathbf{B}^x(a, t^*, \phi) \quad \forall \psi \left( (\mathbf{B}^y(a, t^*, \psi) \wedge \psi \neq \phi) \rightarrow y < x \right)}{\mathbf{K}\left(a, t^*, \mathbf{O}(a, t^*, \text{emergency}, \text{happens}(\text{action}(a^*, \text{land}(\phi)), t^*))\right)} [I_{EP}]
 \end{array}$$

# Technical Preliminaries

- Cognitive Calculi
- Multi-sorted FOL
- Inference Schemata

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

## Inference Schemata

$$\frac{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \rightarrow \mathbf{K}(a, t, \phi)) \quad [I_1]}{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \rightarrow \mathbf{B}(a, t, \phi)) \quad [I_2]}$$

$$\frac{\mathbf{K}(a, t, \phi)}{\phi} \quad [I_4]$$

$$\frac{\mathbf{S}(s, h, t, \phi) \quad [I_{12}]}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} \quad \frac{\mathbf{I}(a, t, happens(action(a^*, \alpha), t')) \quad [I_{13}]}{\mathbf{P}(a, t, happens(action(a^*, \alpha), t'))}$$

$$\frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} \quad [I_{14}]$$

$$\frac{\mathbf{B}^x(a, t^*, \phi) \quad \forall \psi ((\mathbf{B}^y(a, t^*, \psi) \wedge \psi \neq \phi) \rightarrow y < x)}{\mathbf{K}(a, t^*, \mathbf{O}(a, t^*, emergency, happens(action(a^*, land(\phi)), t^*)))} \quad [I_{EP}]$$

# Technical Preliminaries

- Cognitive Calculi
- Multi-sorted FOL
- Inference Schemata
- Strength Factors

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

## Inference Schemata

$$\frac{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \rightarrow \mathbf{K}(a, t, \phi)) \quad [I_1]}{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \rightarrow \mathbf{B}(a, t, \phi)) \quad [I_2]}$$

$$\frac{\mathbf{K}(a, t, \phi)}{\phi} \quad [I_4]$$

$$\frac{\mathbf{S}(s, h, t, \phi) \quad [I_{12}]}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} \quad \frac{\mathbf{I}(a, t, happens(action(a^*, \alpha), t')) \quad [I_{13}]}{\mathbf{P}(a, t, happens(action(a^*, \alpha), t'))}$$

$$\frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} \quad [I_{14}]$$

$$\frac{\mathbf{B}^x(a, t^*, \phi) \quad \forall \psi ((\mathbf{B}^y(a, t^*, \psi) \wedge \psi \neq \phi) \rightarrow y < x)}{\mathbf{K}(a, t^*, \mathbf{O}(a, t^*, emergency, happens(action(a^*, land(\phi)), t^*)))} \quad [I_{EP}]$$

# Technical Preliminaries

- Cognitive Calculi
  - Multi-sorted FOL
  - Inference Schemata
- Strength Factors
  - Reasonableness

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

## Inference Schemata

$$\frac{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \rightarrow \mathbf{K}(a, t, \phi)) \quad [I_1]}{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \rightarrow \mathbf{B}(a, t, \phi)) \quad [I_2]}$$

$$\frac{\mathbf{K}(a, t, \phi)}{\phi} \quad [I_4]$$

$$\frac{\mathbf{S}(s, h, t, \phi) \quad [I_{12}]}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} \quad \frac{\mathbf{I}(a, t, happens(action(a^*, \alpha), t')) \quad [I_{13}]}{\mathbf{P}(a, t, happens(action(a^*, \alpha), t'))}$$

$$\frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} \quad [I_{14}]$$

$$\frac{\mathbf{B}^x(a, t^*, \phi) \quad \forall \psi ((\mathbf{B}^y(a, t^*, \psi) \wedge \psi \neq \phi) \rightarrow y < x)}{\mathbf{K}(a, t^*, \mathbf{O}(a, t^*, emergency, happens(action(a^*, land(\phi)), t^*)))} \quad [I_{EP}]$$

# Technical Preliminaries

- Cognitive Calculi

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

## Inference Schemata

## Domain-Specific Reasonableness

$$\text{Land}(a, t, \phi) \succ_t^a \text{Land}(a, t, \psi) \equiv \mathbf{P}\left(a, t, \left(\begin{array}{l} \text{Reachable}(a, t, \phi) \wedge \neg\text{Reachable}(a, t, \psi) \\ \vee \text{safety}(a, t, \phi) > \text{safety}(a, t, \psi) \end{array}\right)\right) \quad [\succ_t^a \text{-def}]$$

- Strength Factors
  - Reasonableness

# Technical Preliminaries

- Cognitive Calculi

- Multi-sorted FOL

- Inference Schemata

- Strength Factors

- Reasonableness

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

## Inference Schemata

$$\begin{array}{c} \overline{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \rightarrow \mathbf{K}(a, t, \phi))} [I_1] \quad \overline{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \rightarrow \mathbf{B}(a, t, \phi))} [I_2] \\ \frac{\mathbf{K}(a, t, \phi)}{\phi} [I_4] \\ \frac{\mathbf{S}(s, h, t, \phi)}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} [I_{12}] \quad \frac{\mathbf{I}(a, t, happens(action(a^*, \alpha), t'))}{\mathbf{P}(a, t, happens(action(a^*, \alpha), t'))} [I_{13}] \\ \frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} [I_{14}] \\ \frac{\mathbf{B}^x(a, t^*, \phi) \quad \forall \psi ((\mathbf{B}^y(a, t^*, \psi) \wedge \psi \neq \phi) \rightarrow y < x)}{\mathbf{K}(a, t^*, \mathbf{O}(a, t^*, emergency, happens(action(a^*, land(\phi)), t^*)))} [I_{EP}] \end{array}$$

## Domain-Specific Reasonableness

$$\text{Land}(a, t, \phi) \succ_t^a \text{Land}(a, t, \psi) \equiv \mathbf{P}\left(a, t, \left( \begin{array}{l} \text{Reachable}(a, t, \phi) \wedge \neg\text{Reachable}(a, t, \psi) \\ \vee \text{safety}(a, t, \phi) > \text{safety}(a, t, \psi) \end{array} \right) \right) [\succ_t^a \text{-def}]$$

# Technical Preliminaries

- Cognitive Calculi

- Multi-sorted FOL

- Inference Schemata

- Strength Factors

- Reasonableness

- Uncertainty Measures

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

## Inference Schemata

$$\frac{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \rightarrow \mathbf{K}(a, t, \phi)) \quad [I_1]}{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \rightarrow \mathbf{B}(a, t, \phi)) \quad [I_2]}$$

$$\frac{\mathbf{K}(a, t, \phi)}{\phi} \quad [I_4]$$

$$\frac{\mathbf{S}(s, h, t, \phi) \quad [I_{12}]}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} \quad \frac{\mathbf{I}(a, t, happens(action(a^*, \alpha), t')) \quad [I_{13}]}{\mathbf{P}(a, t, happens(action(a^*, \alpha), t'))}$$

$$\frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} \quad [I_{14}]$$

$$\frac{\mathbf{B}^x(a, t^*, \phi) \quad \forall \psi ((\mathbf{B}^y(a, t^*, \psi) \wedge \psi \neq \phi) \rightarrow y < x)}{\mathbf{K}(a, t^*, \mathbf{O}(a, t^*, emergency, happens(action(a^*, land(\phi)), t^*)))} \quad [I_{EP}]$$

## Domain-Specific Reasonableness

$$\text{Land}(a, t, \phi) \succ_t^a \text{Land}(a, t, \psi) \equiv \mathbf{P}\left(a, t, \left( \begin{array}{l} \text{Reachable}(a, t, \phi) \wedge \neg\text{Reachable}(a, t, \psi) \\ \vee \text{safety}(a, t, \phi) > \text{safety}(a, t, \psi) \end{array} \right) \right) \quad [\succ_t^a \text{-def}]$$

# Uncertainty Measures

**More Likely Than Not** Agent  $a$  believes the Air Traffic Controller  $atc$  believes that  $a$  should land at  $\phi$ :

$$\mathbf{B}^1(a, t, \text{Land}(a, t, \phi)) \equiv \mathbf{B}(a, t, \mathbf{B}(atc, t, \text{Land}(a, t, \phi))) \quad [\mathbf{B}^1\text{-def}]$$

**Likely** Agent  $a$  perceives an emergency, and while  $a$  believes the Air Traffic Controller  $atc$  believes  $a$  should land at  $\psi$ ,  $a$  finds it more reasonable to land at  $\phi$ :

$$\mathbf{B}^2(a, t, \text{Land}(a, t, \phi)) \equiv \left( \begin{array}{l} \mathbf{P}(a, t, \text{emergency}) \wedge \mathbf{B}^1(a, t, \text{Land}(a, t, \psi)) \\ \wedge \text{Land}(a, t, \phi) \succ_t^a \text{Land}(a, t, \psi) \end{array} \right) \quad [\mathbf{B}^2\text{-def}]$$

**Beyond Reasonable Doubt** Agent  $a$  perceives an emergency and perceives the safety of landing at  $\phi$  to be higher than some constant threshold  $\gamma$ :

$$\mathbf{B}^3(a, t, \text{Land}(a, t, \phi)) \equiv \mathbf{P}(a, t, \text{emergency}) \wedge \mathbf{P}(a, t, \text{safety}(a, t, \phi) > \gamma) \quad [\mathbf{B}^3\text{-def}]$$

**Evident** Agent  $a$  perceives an emergency, perceives that  $\phi$  meets the safety threshold  $\gamma$ , and believes the Air Traffic Controller  $atc$  believes  $a$  should land at  $\phi$ :

$$\mathbf{B}^4(a, t, \text{Land}(a, t, \phi)) \equiv \mathbf{B}^1(a, t, \text{Land}(a, t, \phi)) \wedge \mathbf{B}^3(a, t, \text{Land}(a, t, \phi)) \quad [\mathbf{B}^4\text{-def}]$$

# Technical Preliminaries

- Cognitive Calculi

- Multi-sorted FOL

- Inference Schemata

- Strength Factors

- Reasonableness

- Uncertainty Measures

## Grammar

$$\phi ::= \begin{cases} q : \text{Formula} \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \forall x : \phi(x) \mid \exists x : \phi(x) \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \\ \mathbf{C}(t, \phi) \mid \mathbf{B}(a, t, \phi) \mid \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg)happens(action(a^*, \alpha), t')) \end{cases}$$

Perceives, Knows, Says, Common-knowledge  
Believes, Desires, Intends, Ought-to

## Inference Schemata

$$\frac{\mathbf{C}(t, \mathbf{P}(a, t, \phi) \rightarrow \mathbf{K}(a, t, \phi)) \quad [I_1]}{\mathbf{C}(t, \mathbf{K}(a, t, \phi) \rightarrow \mathbf{B}(a, t, \phi)) \quad [I_2]}$$

$$\frac{\mathbf{K}(a, t, \phi)}{\phi} \quad [I_4]$$

$$\frac{\mathbf{S}(s, h, t, \phi) \quad [I_{12}]}{\mathbf{B}(h, t, \mathbf{B}(s, t, \phi))} \quad \frac{\mathbf{I}(a, t, happens(action(a^*, \alpha), t')) \quad [I_{13}]}{\mathbf{P}(a, t, happens(action(a^*, \alpha), t'))}$$

$$\frac{\mathbf{B}(a, t, \phi) \quad \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \quad \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} \quad [I_{14}]$$

$$\frac{\mathbf{B}^x(a, t^*, \phi) \quad \forall \psi ((\mathbf{B}^y(a, t^*, \psi) \wedge \psi \neq \phi) \rightarrow y < x)}{\mathbf{K}(a, t^*, \mathbf{O}(a, t^*, emergency, happens(action(a^*, land(\phi)), t^*)))} \quad [I_{EP}]$$

## Domain-Specific Reasonableness

$$\text{Land}(a, t, \phi) \succ_t^a \text{Land}(a, t, \psi) \equiv \mathbf{P}\left(a, t, \left( \begin{array}{l} \text{Reachable}(a, t, \phi) \wedge \neg\text{Reachable}(a, t, \psi) \\ \vee \text{safety}(a, t, \phi) > \text{safety}(a, t, \psi) \end{array} \right) \right) \quad [\succ_t^a \text{-def}]$$

## Uncertainty Measures

**More Likely Than Not** Agent  $a$  believes the Air Traffic Controller  $atc$  believes that  $a$  should land at  $\phi$ :

$$\mathbf{B}^1(a, t, \text{Land}(a, t, \phi)) \equiv \mathbf{B}(a, t, \mathbf{B}(atc, t, \text{Land}(a, t, \phi))) \quad [\mathbf{B}^1\text{-def}]$$

**Likely** Agent  $a$  perceives an emergency, and while  $a$  believes the Air Traffic Controller  $atc$  believes  $a$  should land at  $\psi$ ,  $a$  finds it more reasonable to land at  $\phi$ :

$$\mathbf{B}^2(a, t, \text{Land}(a, t, \phi)) \equiv \left( \begin{array}{l} \mathbf{P}(a, t, \text{emergency}) \wedge \mathbf{B}^1(a, t, \text{Land}(a, t, \psi)) \\ \wedge \text{Land}(a, t, \phi) \succ_t^a \text{Land}(a, t, \psi) \end{array} \right) \quad [\mathbf{B}^2\text{-def}]$$

**Beyond Reasonable Doubt** Agent  $a$  perceives an emergency and perceives the safety of landing at  $\phi$  to be higher than some constant threshold  $\gamma$ :

$$\mathbf{B}^3(a, t, \text{Land}(a, t, \phi)) \equiv \mathbf{P}(a, t, \text{emergency}) \wedge \mathbf{P}(a, t, \text{safety}(a, t, \phi) > \gamma) \quad [\mathbf{B}^3\text{-def}]$$

**Evident** Agent  $a$  perceives an emergency, perceives that  $\phi$  meets the safety threshold  $\gamma$ , and believes the Air Traffic Controller  $atc$  believes  $a$  should land at  $\phi$ :

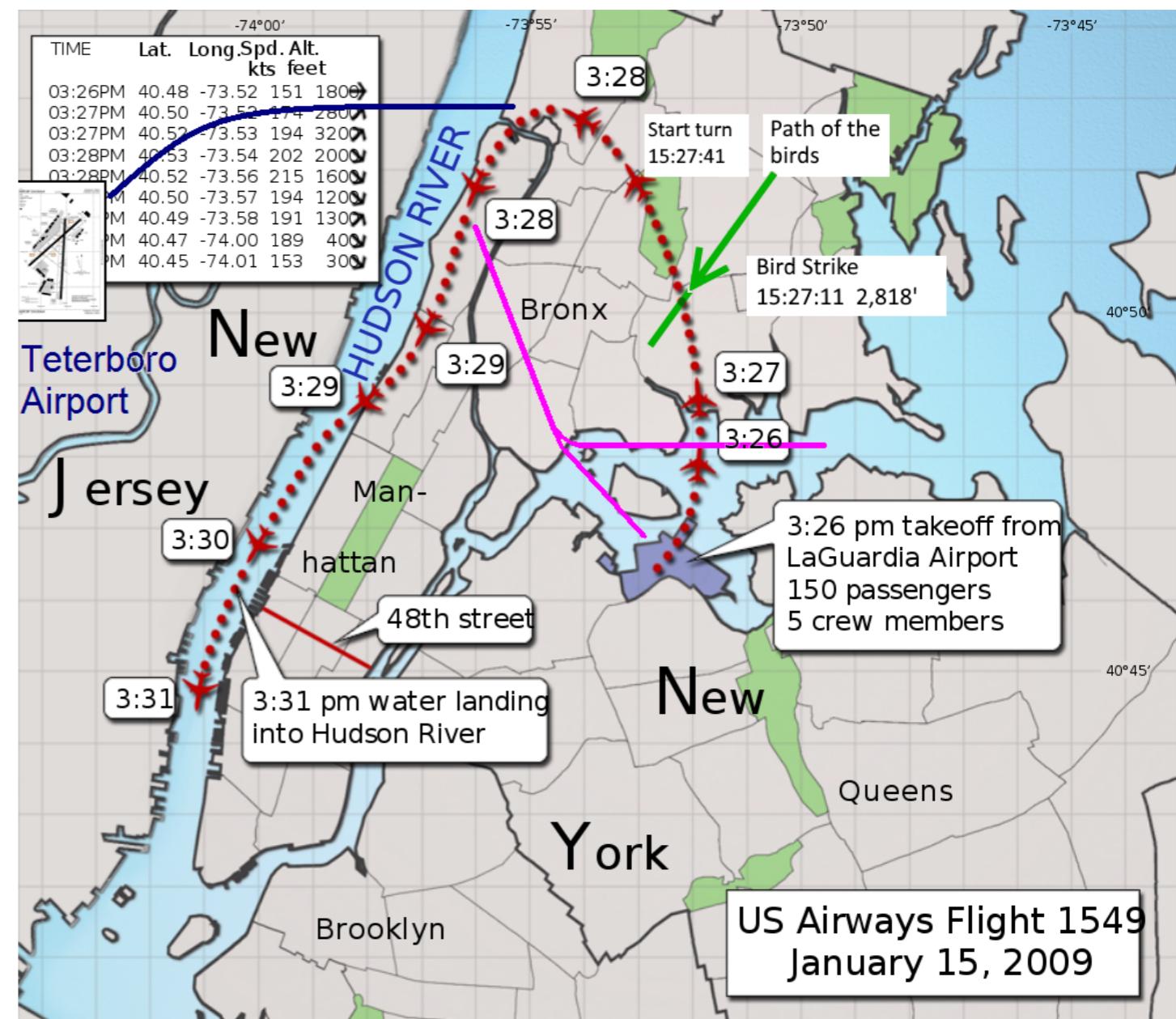
$$\mathbf{B}^4(a, t, \text{Land}(a, t, \phi)) \equiv \mathbf{B}^1(a, t, \text{Land}(a, t, \phi)) \wedge \mathbf{B}^3(a, t, \text{Land}(a, t, \phi)) \quad [\mathbf{B}^4\text{-def}]$$

# Miracle on the Hudson

- Famous water landing
- Necessitated by a twin engine failure when the plane flew into a large flock of Canada geese



[https://upload.wikimedia.org/wikipedia/commons/8/8f/US\\_Airways\\_Flight\\_1549\\_%28N106US%29\\_after\\_crashing\\_into\\_the\\_Hudson\\_River\\_%28crop\\_2%29.jpg](https://upload.wikimedia.org/wikipedia/commons/8/8f/US_Airways_Flight_1549_%28N106US%29_after_crashing_into_the_Hudson_River_%28crop_2%29.jpg)



[https://upload.wikimedia.org/wikipedia/commons/0/03/Flight\\_1549-OptionsNotTaken.PNG](https://upload.wikimedia.org/wikipedia/commons/0/03/Flight_1549-OptionsNotTaken.PNG)

TEB



LGA



TEB

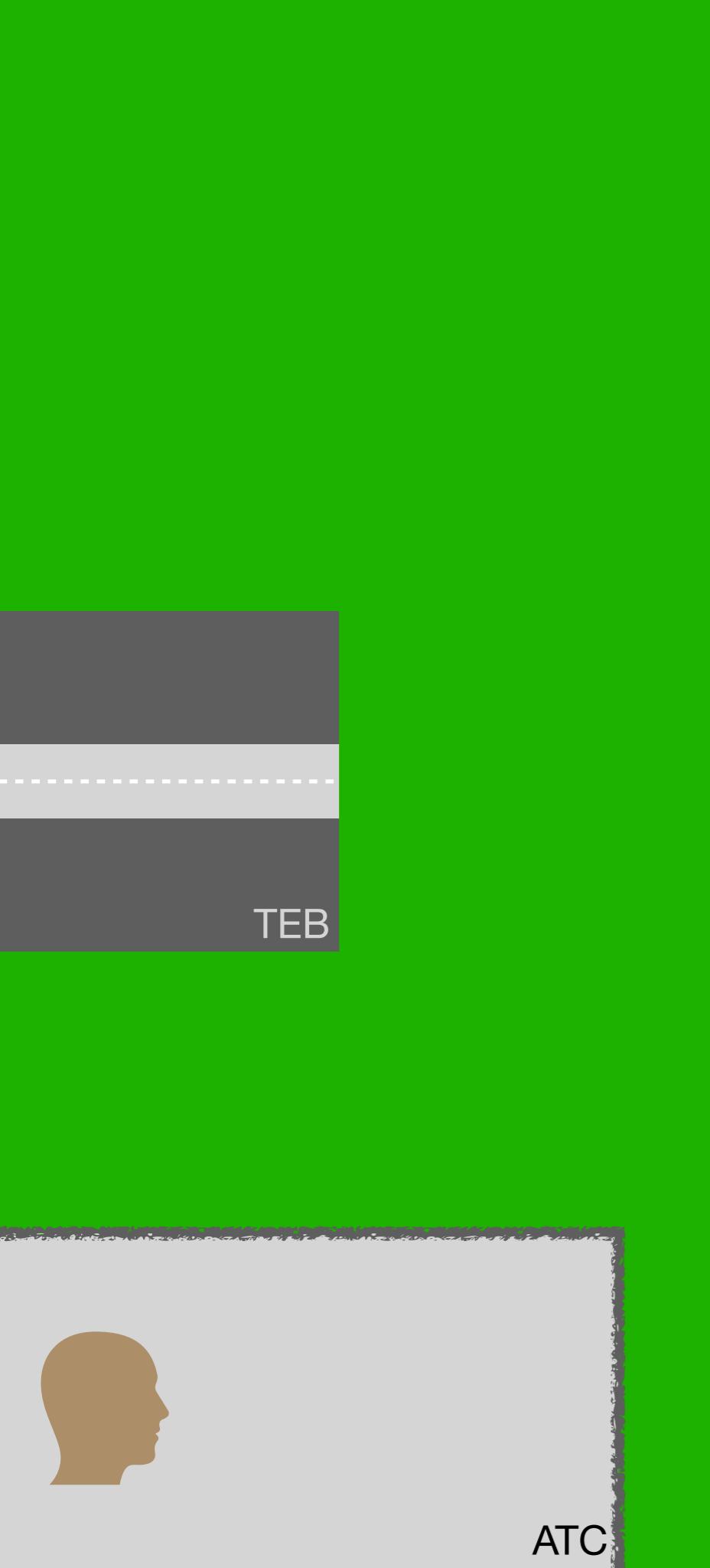
LGA

TEB

“Birds”



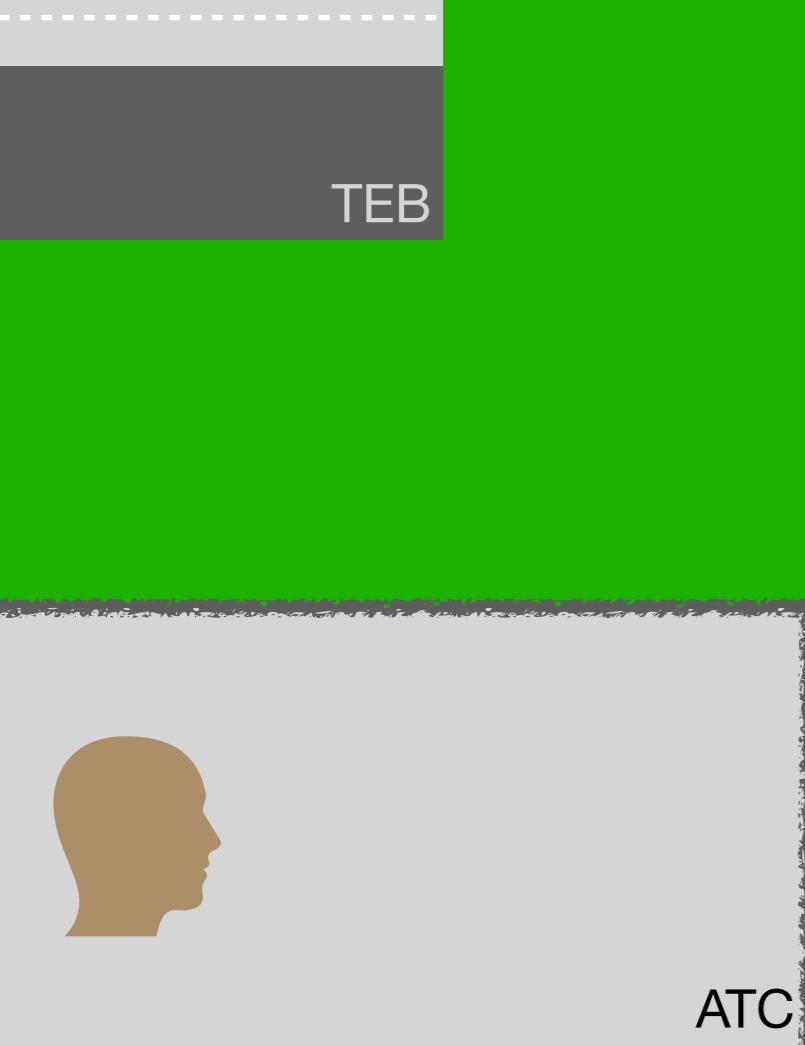
LGA



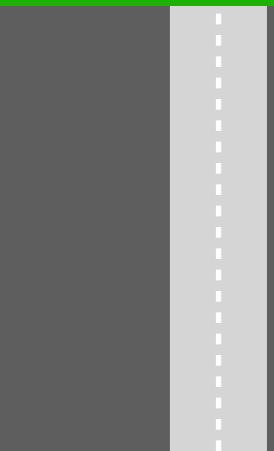
TEB



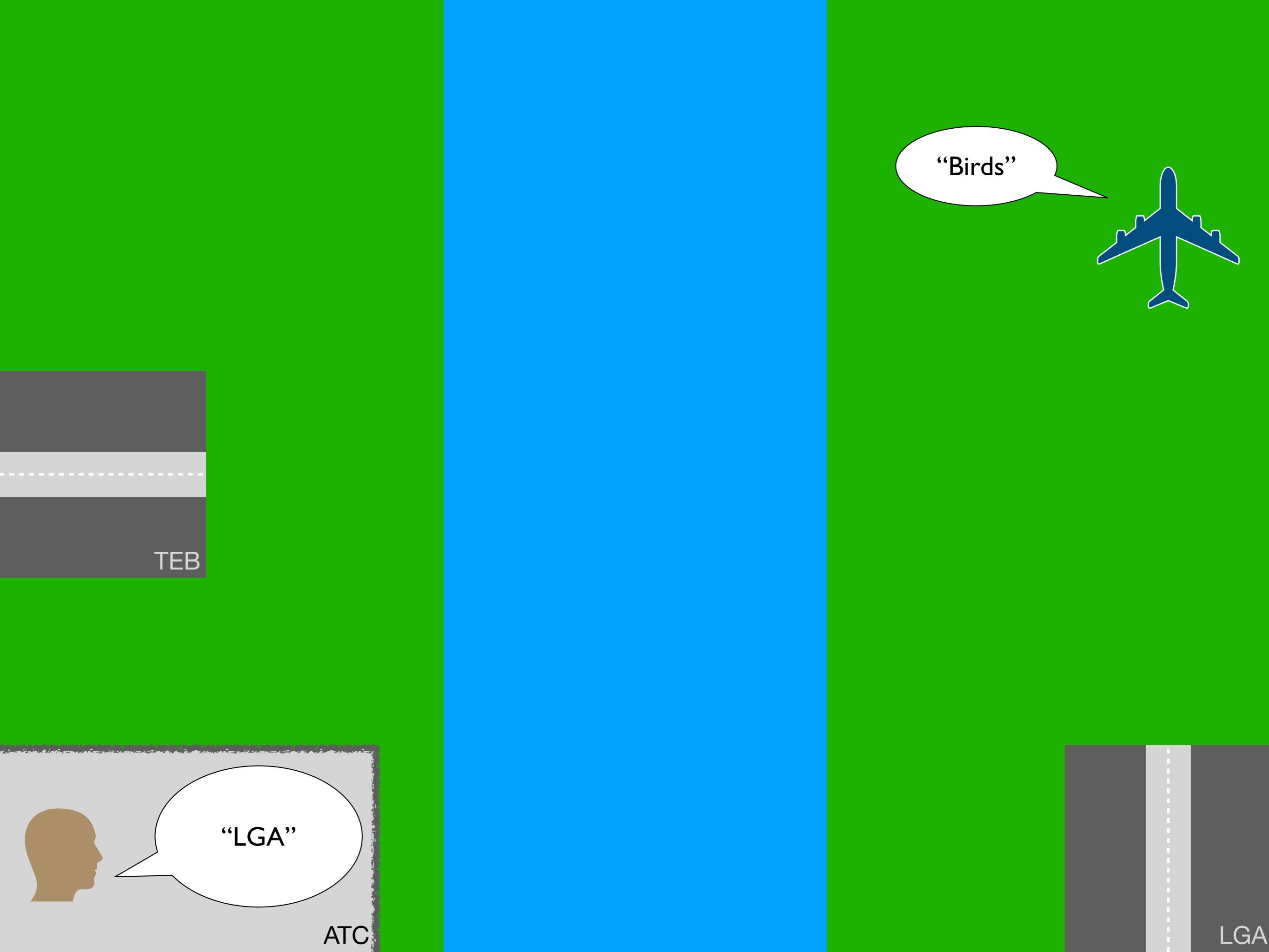
ATC



“Birds”



LGA



$\mathbf{S}(atc, capt, t_1, \text{Land}(capt, t_1, lga_{13}))$

$\therefore \mathbf{B}(capt, t_1, \mathbf{B}(atc, t_1, \text{Land}(capt, t_1, lga_{13}))) [I_{12}] \checkmark$

$\therefore \mathbf{B}^1(capt, t_1, \text{Land}(capt, t_1, lga_{13})) [\mathbf{B}^1\text{-def}] \checkmark$

TEB

“LGA”

ATC

“Birds”



LGA



TEB

“LGA”



ATC

LGA

“LGA  
unreachable”



TEB

“LGA”

ATC

LGA

This diagram illustrates a communication exchange between Air Traffic Control (ATC) and Terminal Equipment Board (TEB) regarding the status of LaGuardia Airport (LGA). The communication is shown through speech bubbles and a central blue vertical bar.

The diagram is divided into two main vertical sections: a grey section on the left and a green section on the right, separated by a blue central bar.

**ATC (Left):**

- A brown head icon represents the ATC operator.
- A speech bubble contains the text "LGA".
- The word "ATC" is written below the central bar.

**TEB (Right):**

- A blue airplane icon represents the aircraft.
- Two speech bubbles are present:
  - The top bubble contains the text "LGA unreachable".
  - The bottom bubble contains the text "TEB".
- The word "TEB" is written below the central bar.

The blue central bar serves as a visual connector between the two sections, representing the communication link between ATC and TEB.

"LGA"  
"LGA unreachable"  
"TEB"

$\therefore \text{Land}(\text{capt}, t_2, \text{teb}) \succ_{t_2}^{\text{capt}} \text{Land}(\text{capt}, t_2, \text{lga}_{13})$  [ $\succ_t^a$  -def] ✓  
 $\therefore \mathbf{B}^2(\text{capt}, t_2, \text{Land}(\text{capt}, t_2, \text{teb}))$  [ $\mathbf{B}^2$ -def] ✓

“LGA  
unreachable”



“TEB”

TEB

“LGA”



ATC

LGA



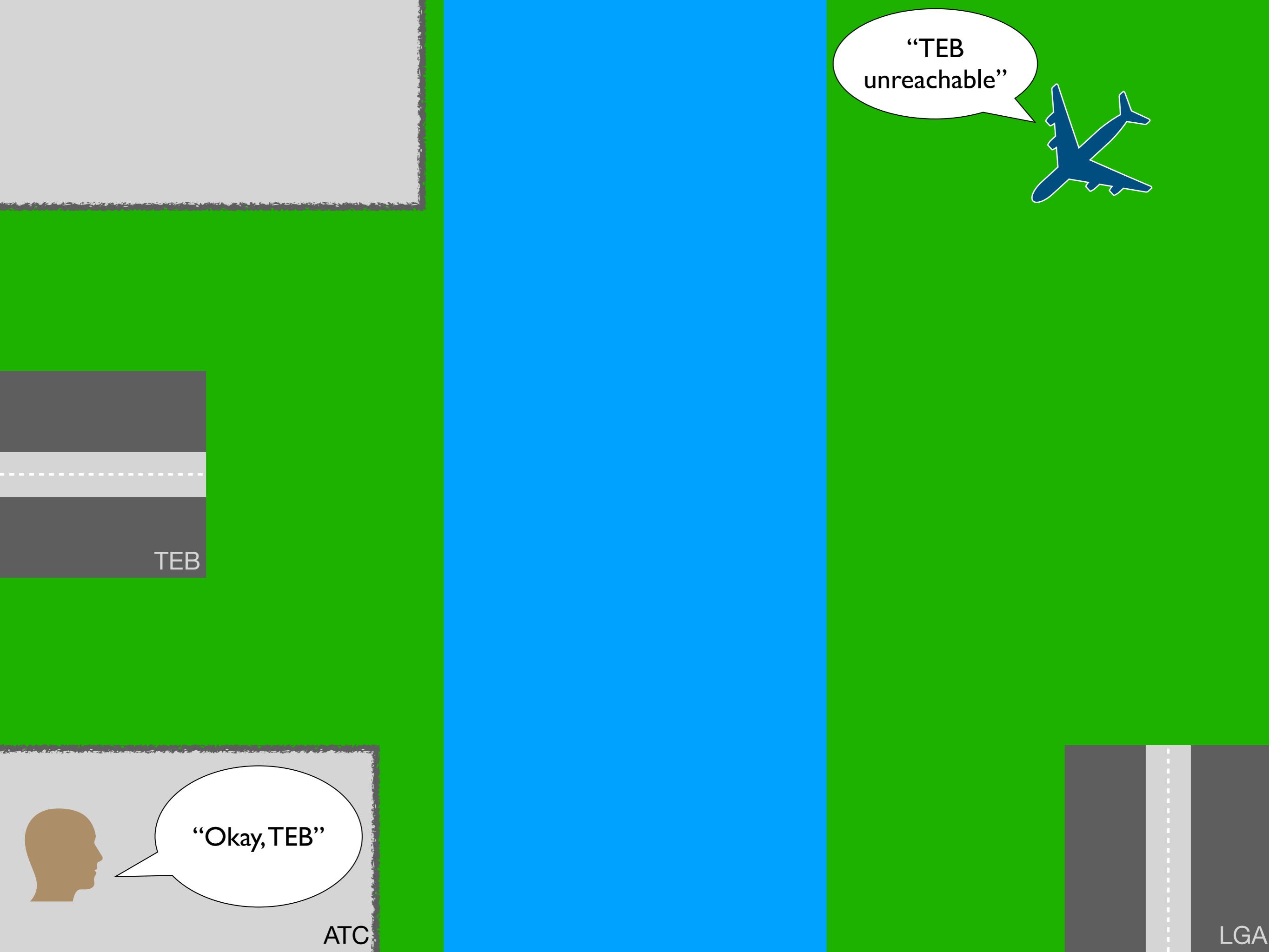
TEB

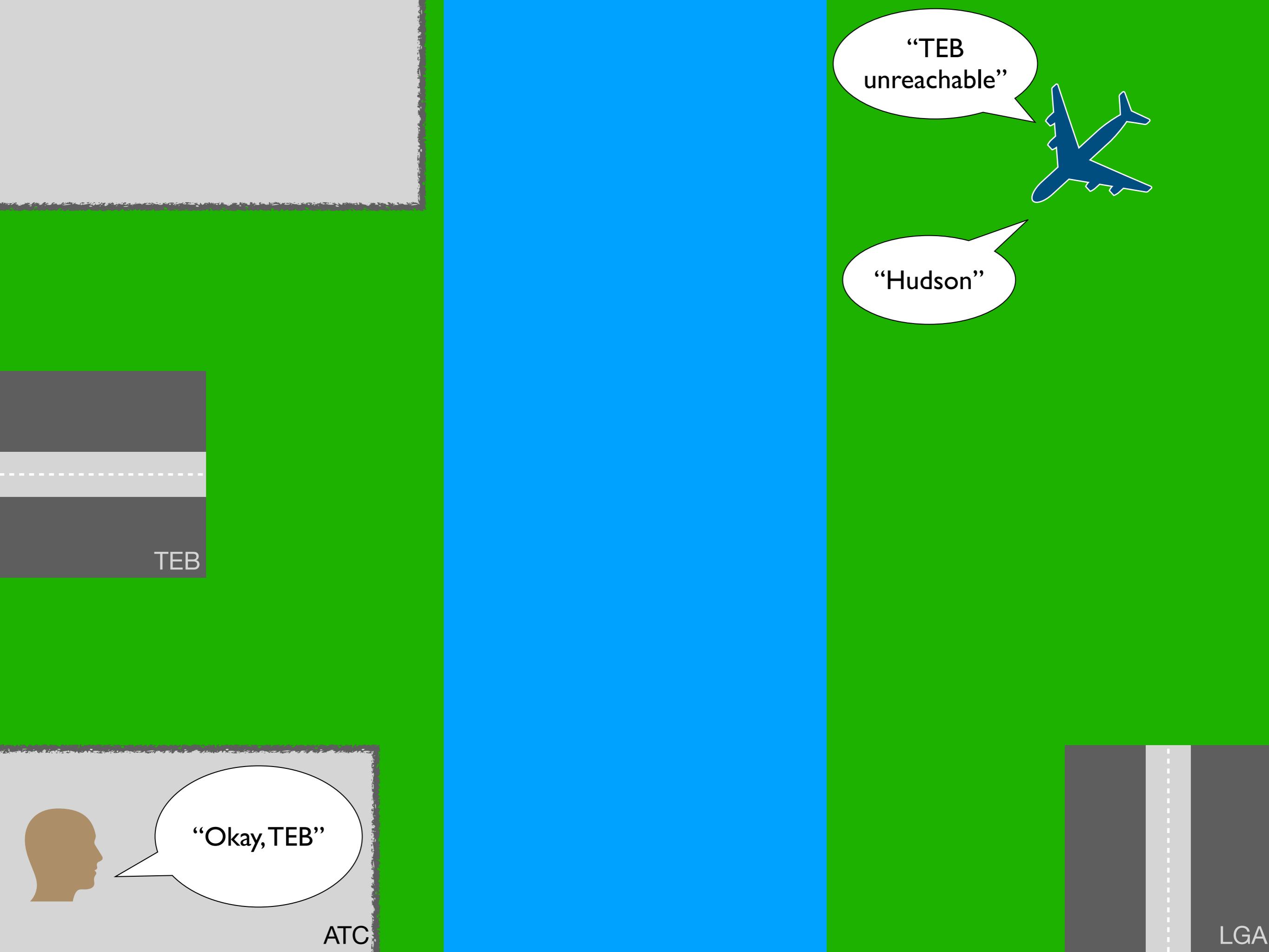
"Okay, TEB"



ATC

LGA





$S(atc, capt, t_3, \text{Land}(capt, t_3, teb_1))$

$\therefore B(capt, t_3, B(atc, t_3, \text{Land}(capt, t_3, teb_1))) [I_{12}] \checkmark$

$\therefore B^1(capt, t_3, \text{Land}(capt, t_3, teb_1)) [B^1\text{-def}] \checkmark$

$\therefore \text{Land}(capt, t_3, hud) \succ_{t_3}^{capt} \text{Land}(capt, t_3, teb_1) [\succ_t^a \text{-def}] \checkmark$

$\therefore B^2(capt, t_3, \text{Land}(capt, t_3, hud)) [B^2\text{-def}] \checkmark$

TEB

“Okay, TEB”

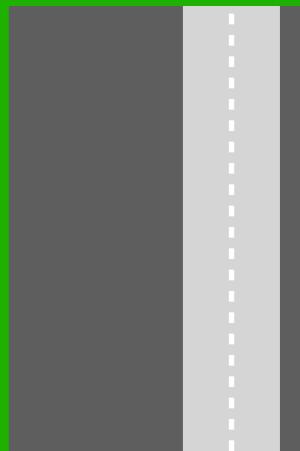


ATC

“TEB  
unreachable”



“Hudson”



LGA

$\mathbf{S}(atc, capt, t_3, \text{Land}(capt, t_3, teb_1))$

$\therefore \mathbf{B}(capt, t_3, \mathbf{B}(atc, t_3, \text{Land}(capt, t_3, teb_1))) [I_{12}] \checkmark$

$\therefore \mathbf{B}^1(capt, t_3, \text{Land}(capt, t_3, teb_1)) [\mathbf{B}^1\text{-def}] \checkmark$

$\therefore \text{Land}(capt, t_3, hud) \succ_{t_3}^{capt} \text{Land}(capt, t_3, teb_1) [\succ_t^a \text{-def}] \checkmark$

$\therefore \mathbf{B}^2(capt, t_3, \text{Land}(capt, t_3, hud)) [\mathbf{B}^2\text{-def}] \checkmark$

TEB

“Okay, TEB”

ATC

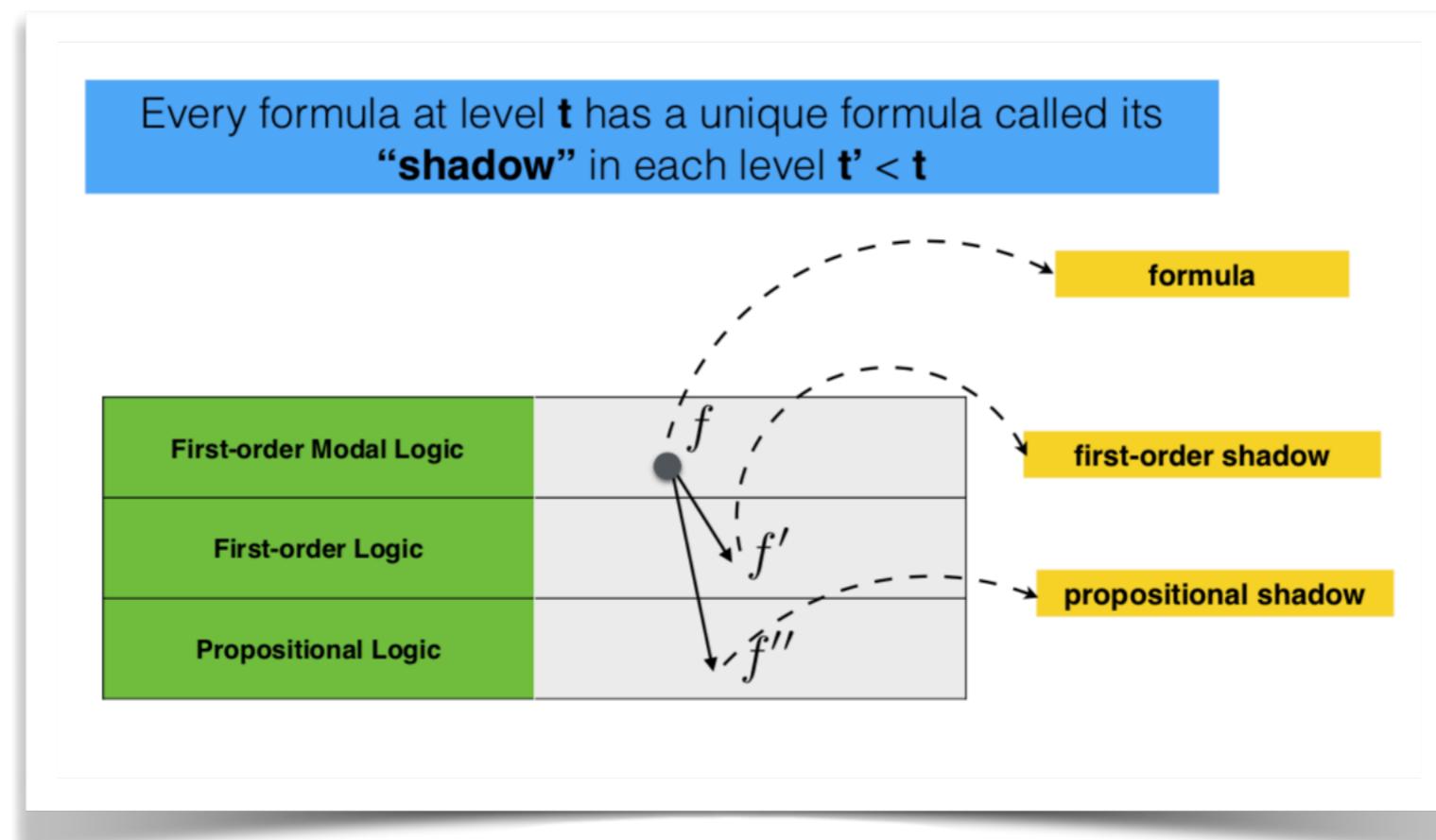


LGA

# ShadowProver

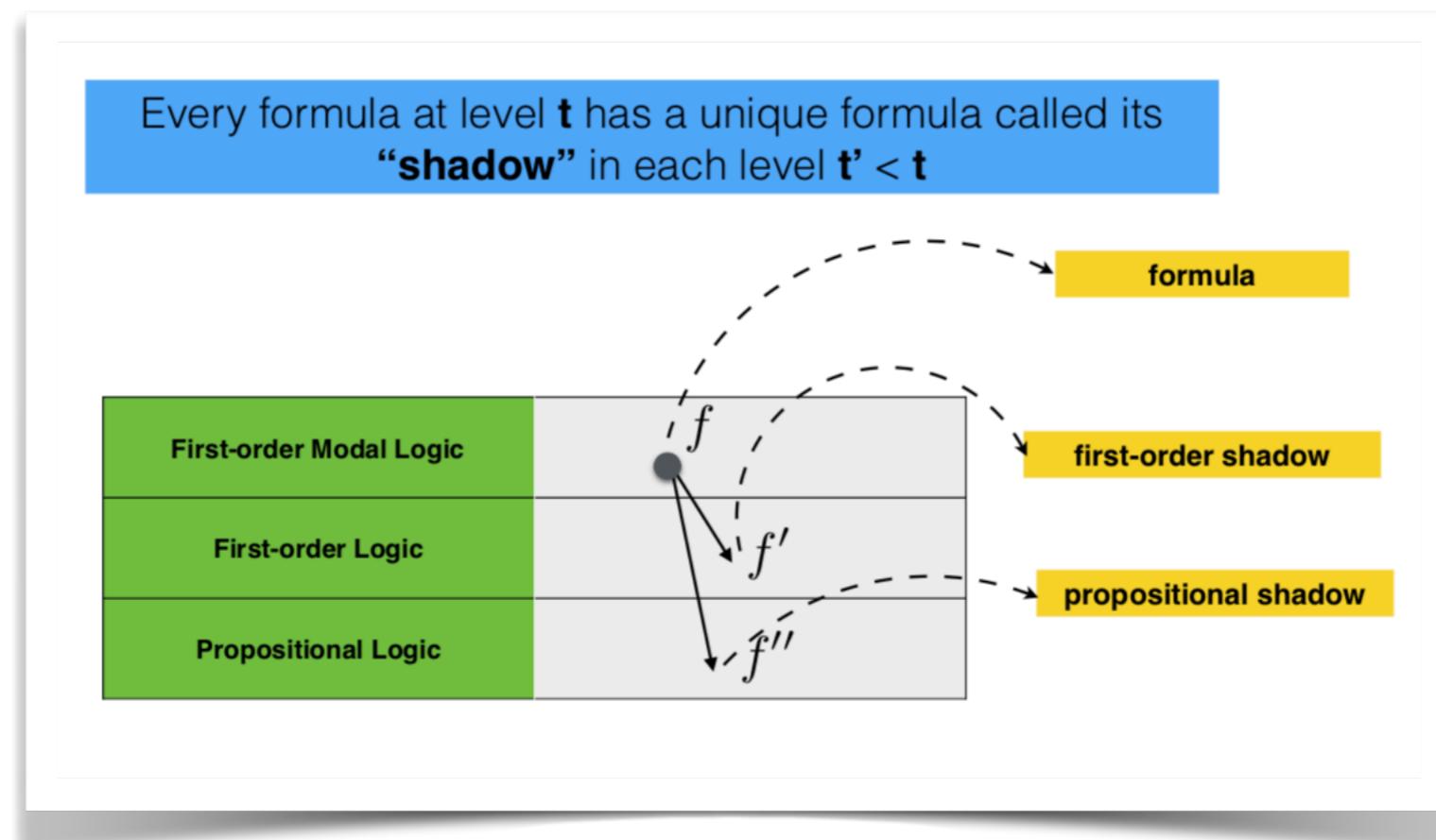
# ShadowProver

- Uses novel “shadowing” technique to find proofs of statements in cognitive calculi



# ShadowProver

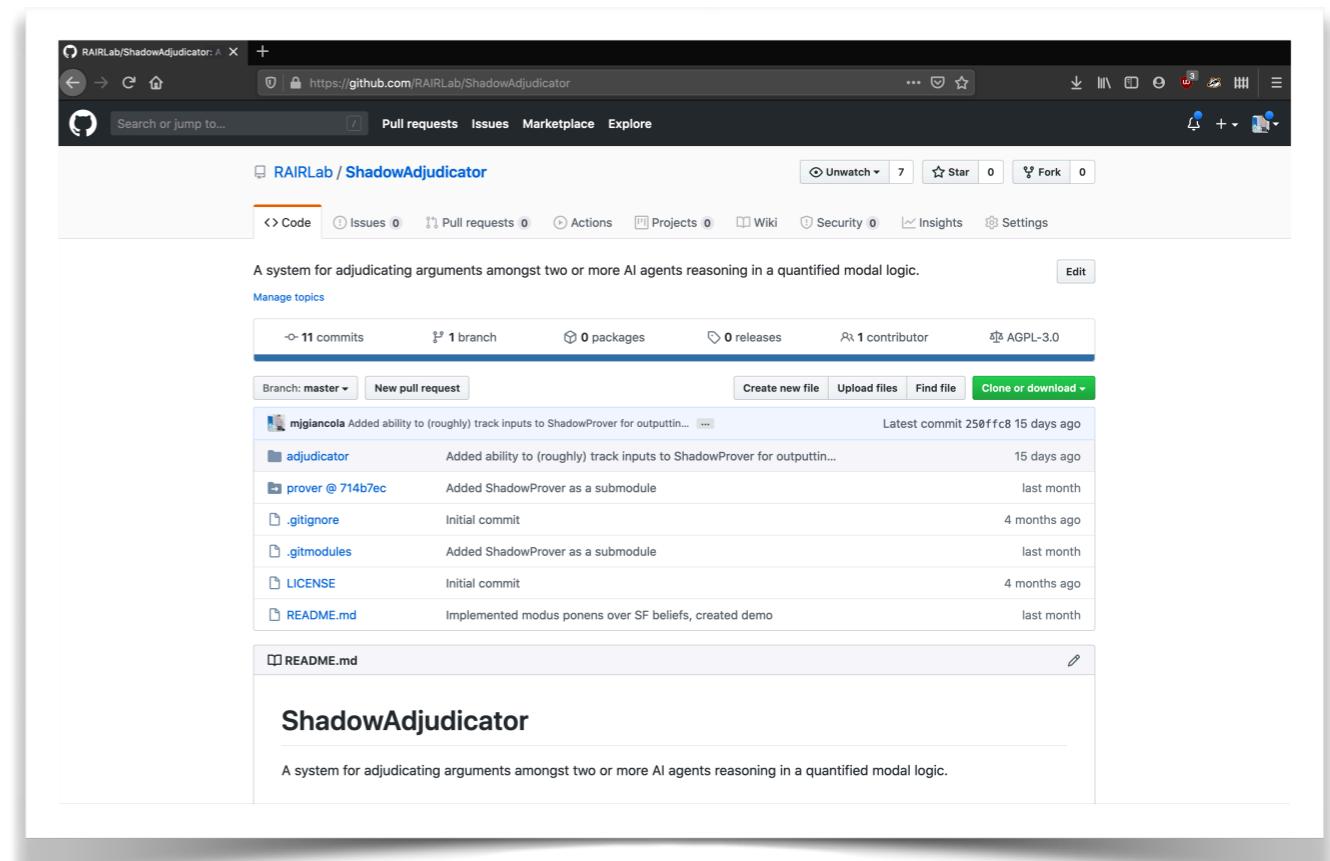
- Uses novel “shadowing” technique to find proofs of statements in cognitive calculi
- However, is not capable of generating inductive arguments



# ShadowAdjudicator

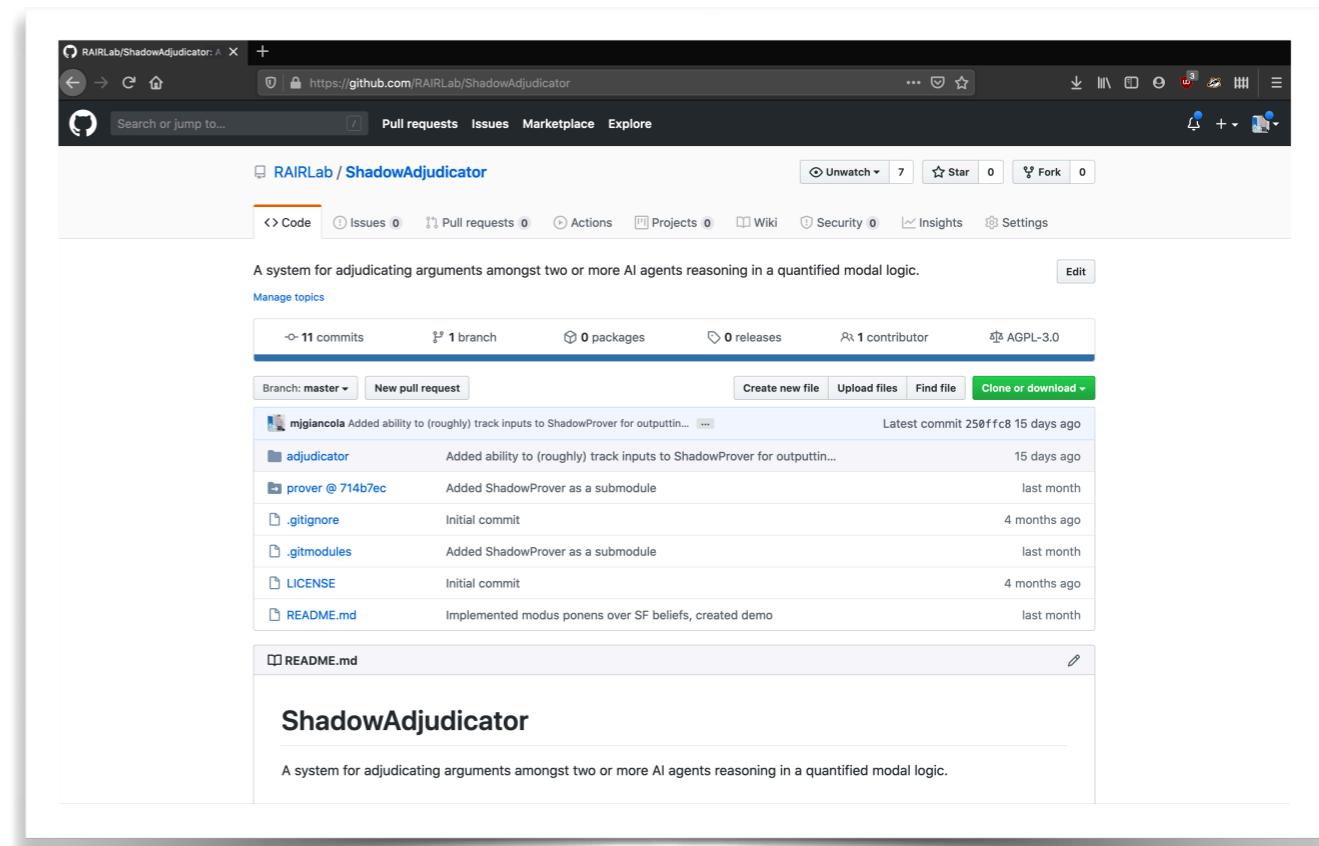
# ShadowAdjudicator

- A nascent automated reasoner for generating and adjudicating arguments



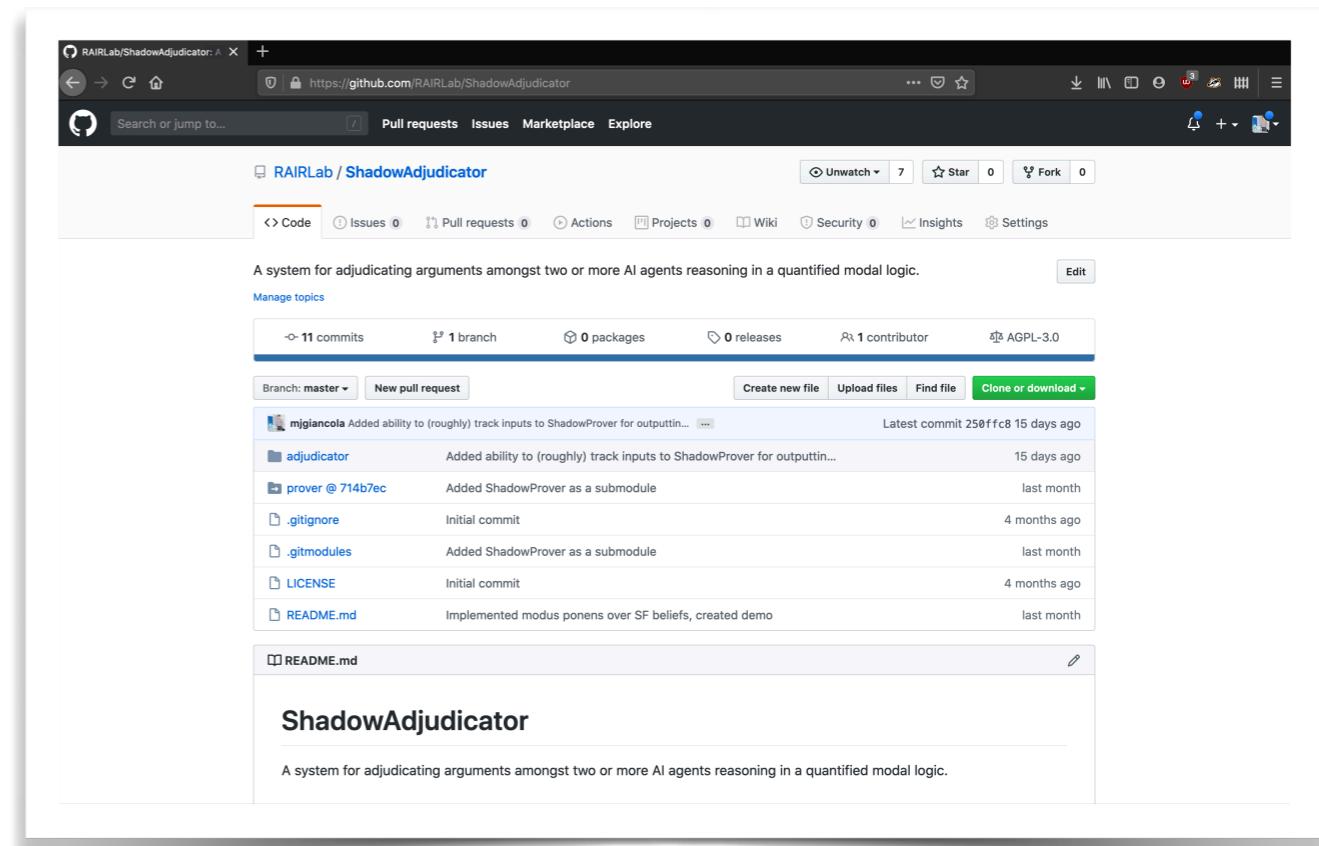
# ShadowAdjudicator

- A nascent automated reasoner for generating and adjudicating arguments
- Builds upon ShadowProver



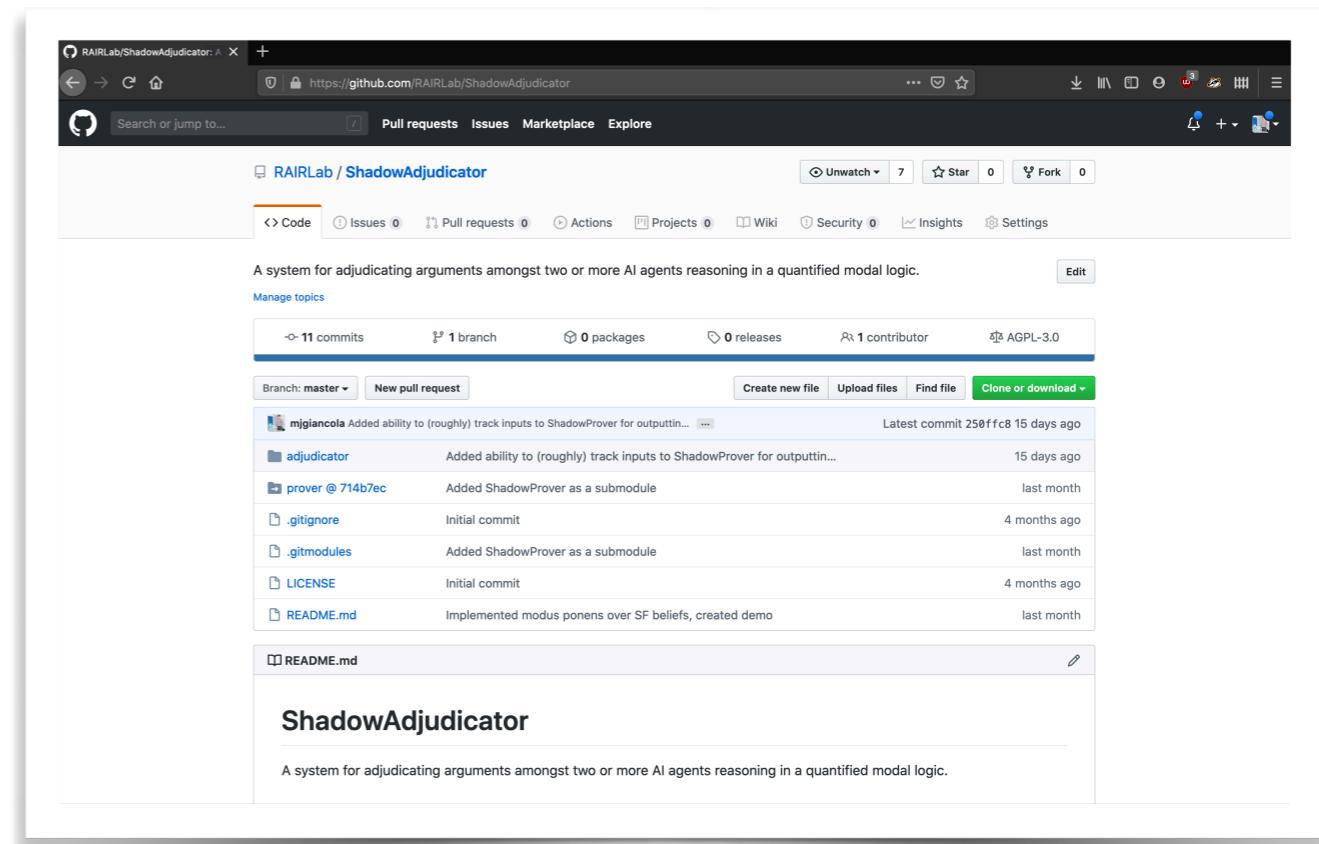
# ShadowAdjudicator

- A nascent automated reasoner for generating and adjudicating arguments
- Builds upon ShadowProver
  - Uses ShadowProver for sub-proofs of modal/FOL/PL formulae

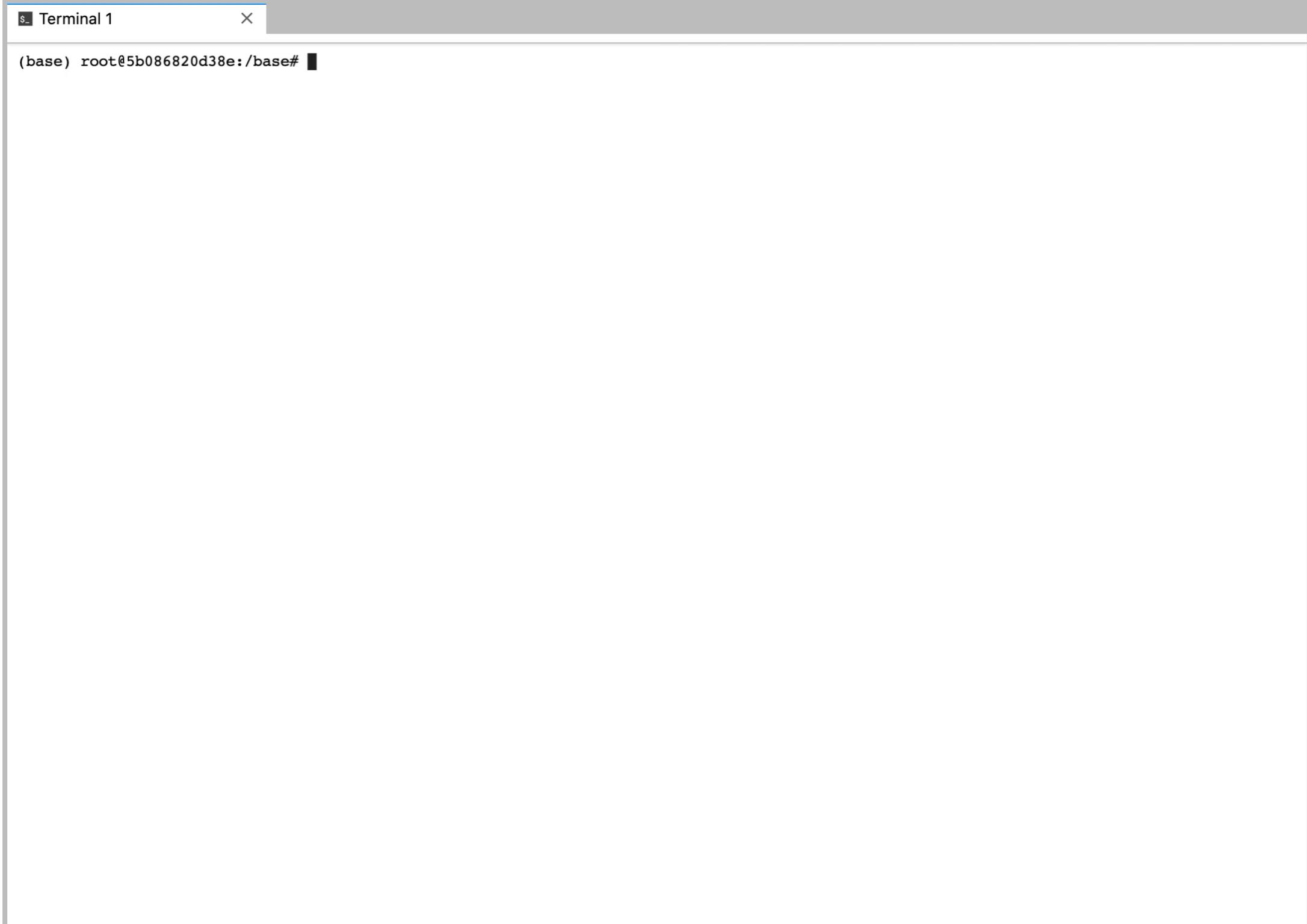


# ShadowAdjudicator

- A nascent automated reasoner for generating and adjudicating arguments
- Builds upon ShadowProver
  - Uses ShadowProver for sub-proofs of modal/FOL/PL formulae
  - Implements an algorithm and inference schemata for generating arguments with strength factors



# Simulation

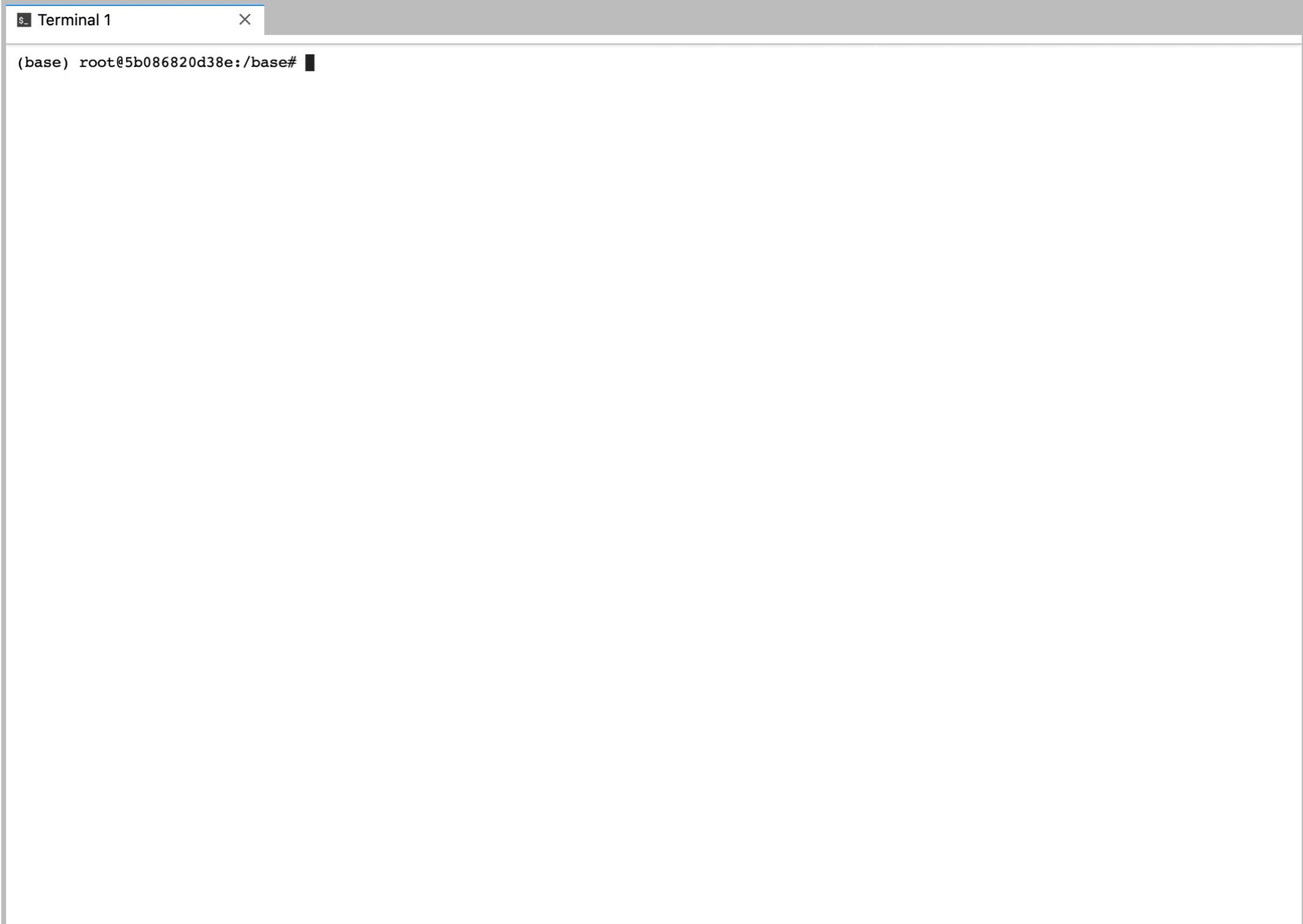


A screenshot of a terminal window titled "Terminal 1". The title bar is grey with white text. The main area of the terminal is white and contains the following text:

```
$ Terminal 1 X  
(base) root@5b086820d38e:/base#
```

The terminal window has a thin grey border.

# Simulation



A screenshot of a terminal window titled "Terminal 1". The title bar is grey with white text. The main area of the terminal is white and contains the following text:

```
$ Terminal 1 X  
(base) root@5b086820d38e:/base#
```

The terminal window has a thin grey border.

# Questions?

# ShadowAdjudicator Arguments

```

1 (base) root@df267ad26288:/base# python demos/2020_ICRES_Ethical_Reasoning/demo_MotH.py
2 Running ShadowProver warmup...
3 Done.
4
5 Generating a solution to Sub-Argument 1...
6 Calling ShadowProver...
7 ShadowProver Done.
8 Argument found in 2.5495572090148926 seconds.
9 PROOF OF: (Believes!1 capt t1 (Land capt t1 lga13))
10 Applied 'Level-1 Definition' to: (Believes! capt t1 (Believes! atc t1 (Land capt t1 lga13)))
11 PROOF OF: (Believes! capt t1 (Believes! atc t1 (Land capt t1 lga13)))
12 Proved via ShadowProver:
13 -->(AssumptionsNowContainsGoal
14 --> Givens:
15 -->((Perceives! capt t1 (Says! atc t1 (Land capt t1 lga13)))
16 --> (GIVEN[]))
17 -->
18 -->((Knows! capt t1 (Says! atc t1 (Land capt t1 lga13)))
19 --> (Perception to knowledge (Perceives! capt t1 (Says! atc t1 (Land capt t1 lga13)))
20 --> [(GIVEN[])])
21 -->
22 -->((Believes! capt t1 (Says! atc t1 (Land capt t1 lga13)))
23 --> InferenceJustification{base=[(Knows! capt t1 (Says! atc t1 (Land capt t1 lga13))), message='DR5']}
24 -->(Believes! capt t1 (Believes! atc t1 (Land capt t1 lga13)))
25 --> Goals:
26 -->(Believes! capt t1 (Believes! atc t1 (Land capt t1 lga13)))
27 PROOF OF: (Perceives! capt t1 (Says! atc t1 (Land capt t1 lga13)))
28 GIVEN
29
30
31 Generating a solution to Sub-Argument 2...
32 Calling ShadowProver...
33 ShadowProver Done.
34 Calling ShadowProver...
35 ShadowProver Done.
36 Argument found in 4.294346809387207 seconds.
37 PROOF OF: (Believes!2 capt t2 (Land capt t2 teb))
38 Applied 'Level-2 Definition' to: (and (Perceives! capt t2 emergency) (Believes!1 capt t2 (Land capt t2 lga13)) (R capt t2 (Land capt t2 teb) (Land capt t2 lga13)))
39 PROOF OF: (and (Perceives! capt t2 emergency) (Believes!1 capt t2 (Land capt t2 lga13)) (R capt t2 (Land capt t2 teb) (Land capt t2 lga13)))
40 Proved via ShadowProver [and a sub-argument of (Believes!1 capt t2 (Land capt t2 lga13)); (R capt t2 (Land capt t2 teb) (Land capt t2 lga13))]:
41 -->(:FOLFromSnark
42 --> Givens:
43 -->((R capt t2 (Land capt t2 teb) (Land capt t2 lga13))
44 --> (GIVEN[]))
45 -->|_Perceives!_capt_t2__Reachable_capt_t2_teb__|
46 -->|_Believes!_capt_t2__Land_capt_t2_lga13__|
47 -->|_Perceives!_capt_t2__not__Reachable_capt_t2_lga13__|
48 -->|_Perceives!_capt_t2_emergency_|)
49 --> Goals:
50 -->|_Perceives!_capt_t2_emergency_|)
51 PROOF OF: (R capt t2 (Land capt t2 teb) (Land capt t2 lga13))
52 Applied 'Reasonableness Definition' to: (Perceives! capt t2 (Reachable capt t2 teb)); (Perceives! capt t2 (not (Reachable capt t2 lga13)))
53 PROOF OF: (Perceives! capt t2 (Reachable capt t2 teb))
54 GIVEN
55 PROOF OF: (Perceives! capt t2 (not (Reachable capt t2 lga13)))
56 GIVEN
57 PROOF OF: (Believes!1 capt t2 (Land capt t2 lga13))
58 GIVEN
59
60
61 Generating a solution to Sub-Argument 3...
62 Calling ShadowProver...

```

```

63 ShadowProver Done.
64 Argument found in 1.9968616962432861 seconds.
65 PROOF OF: (Believes!1 capt t3 (Land capt t3 teb1))
66 Applied 'Level-1 Definition' to: (Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
67 PROOF OF: (Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
68 Proved via ShadowProver:
69 ---(AssumptionsNowContainsGoal
70 --- Givens:
71 --->((Perceives! capt t3 (Reachable capt t3 hud))
72 ---> (GIVEN[]))
73 --->
74 --->((Knows! capt t3 (Reachable capt t3 hud))
75 ---> (Perception to knowledge (Perceives! capt t3 (Reachable capt t3 hud)))
76 ---> [(GIVEN[])])
77 --->
78 --->((Believes! capt t3 (Reachable capt t3 hud))
79 ---> InferenceJustification{base=[(Knows! capt t3 (Reachable capt t3 hud))], message='DR5'})
80 --->
81 --->((Perceives! capt t3 (not (Reachable capt t3 teb1)))
82 ---> (GIVEN[]))
83 --->
84 --->((Knows! capt t3 (not (Reachable capt t3 teb1)))
85 ---> (Perception to knowledge (Perceives! capt t3 (not (Reachable capt t3 teb1)))
86 ---> [(GIVEN[])])
87 --->
88 --->((Believes! capt t3 (not (Reachable capt t3 teb1)))
89 ---> InferenceJustification{base=[(Knows! capt t3 (not (Reachable capt t3 teb1)))]}, message='DR5')
90 --->
91 --->((Perceives! capt t3 emergency)
92 ---> (GIVEN[]))
93 --->
94 --->((Knows! capt t3 emergency)
95 ---> (Perception to knowledge (Perceives! capt t3 emergency)
96 ---> [(GIVEN[])])
97 --->
98 --->((Believes! capt t3 emergency)
99 ---> InferenceJustification{base=[(Knows! capt t3 emergency)], message='DR5'})
100 --->
101 --->((Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1)))
102 ---> (GIVEN[]))
103 --->
104 --->((Knows! capt t3 (Says! atc t3 (Land capt t3 teb1)))
105 ---> (Perception to knowledge (Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1)))
106 ---> [(GIVEN[])])
107 --->
108 --->((Believes! capt t3 (Says! atc t3 (Land capt t3 teb1)))
109 ---> InferenceJustification{base=[(Knows! capt t3 (Says! atc t3 (Land capt t3 teb1)))]}, message='DR5')
110 --->(Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
111 --->
112 --->((R capt t3 (Land capt t3 hud) (Land capt t3 lga13))
113 ---> (GIVEN[]))
114 --->
115 --->((Perceives! capt t3 (not (Reachable capt t3 lga13)))
116 ---> (GIVEN[]))
117 --->
118 --->((Knows! capt t3 (not (Reachable capt t3 lga13)))
119 ---> (Perception to knowledge (Perceives! capt t3 (not (Reachable capt t3 lga13)))
120 ---> [(GIVEN[])])
121 --->
122 --->((Believes! capt t3 (not (Reachable capt t3 lga13)))
123 ---> InferenceJustification{base=[(Knows! capt t3 (not (Reachable capt t3 lga13)))]}, message='DR5')
124 --->

```

```

125 -->((R capt t3 (Land capt t3 hud) (Land capt t3 teb1))
126   -->  (GIVEN[]))
127   --> Goals:
128   --> (Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1))))
129     PROOF OF: (Perceives! capt t3 emergency)
130     GIVEN
131     PROOF OF: (Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1)))
132     GIVEN
133     PROOF OF: (Perceives! capt t3 (not (Reachable capt t3 lga13)))
134     GIVEN
135     PROOF OF: (Perceives! capt t3 (not (Reachable capt t3 teb1)))
136     GIVEN
137     PROOF OF: (Perceives! capt t3 (Reachable capt t3 hud))
138     GIVEN
139     PROOF OF: (R capt t3 (Land capt t3 hud) (Land capt t3 lga13))
140     Applied 'Reasonableness Definition' to: (Perceives! capt t3 (Reachable capt t3 hud)); (Perceives! capt t3 (not (Reachable capt t3 lga13)))
141     PROOF OF: (R capt t3 (Land capt t3 hud) (Land capt t3 teb1))
142     Applied 'Reasonableness Definition' to: (Perceives! capt t3 (Reachable capt t3 hud)); (Perceives! capt t3 (not (Reachable capt t3 teb1)))
143
144 Calling ShadowProver...
145 ShadowProver Done.
146 Calling ShadowProver...
147 ShadowProver Done.
148 Calling ShadowProver...
149 ShadowProver Done.
150 Argument found in 11.011409044265747 seconds.
151 PROOF OF: (Believes!2 capt t3 (Land capt t3 hud))
152 Applied 'Level-2 Definition' to: (and (Perceives! capt t3 emergency) (Believes!1 capt t3 (Land capt t3 teb1)) (R capt t3 (Land capt t3 hud) (Land capt t3 teb1)))
153   PROOF OF: (and (Perceives! capt t3 emergency) (Believes!1 capt t3 (Land capt t3 teb1)) (R capt t3 (Land capt t3 hud) (Land capt t3 teb1)))
154   Proved via ShadowProver [and a sub-argument of (Believes!1 capt t3 (Land capt t3 teb1)); (R capt t3 (Land capt t3 hud) (Land capt t3 teb1))]:
155   -->(:FOLFromSnark
156     --> Givens:
157     -->(|_Believes!_capt_t3__Believes!_atc_t3__Land_capt_t3_teb1___|
158     -->|_Believes!_capt_t3__Land_capt_t3_teb1__|
159     -->|_Perceives!_capt_t3_emergency_|_
160     -->|_Perceives!_capt_t3__not__Reachable_capt_t3_lga13___|
161     -->|_Perceives!_capt_t3__Reachable_capt_t3_hud__|
162     -->|_Perceives!_capt_t3__not__Reachable_capt_t3_teb1___|
163     -->|_Perceives!_capt_t3__Says!_atc_t3__Land_capt_t3_teb1___|
164   -->
165   -->((R capt t3 (Land capt t3 hud) (Land capt t3 lga13))
166     -->  (GIVEN[]))
167   -->
168   -->((R capt t3 (Land capt t3 hud) (Land capt t3 teb1))
169     -->  (GIVEN[]))
170   --> Goals:
171   -->|_Perceives!_capt_t3_emergency_|_
172     PROOF OF: (R capt t3 (Land capt t3 hud) (Land capt t3 lga13))
173     Applied 'Reasonableness Definition' to: (Perceives! capt t3 (Reachable capt t3 hud)); (Perceives! capt t3 (not (Reachable capt t3 lga13)))
174       PROOF OF: (Perceives! capt t3 (Reachable capt t3 hud))
175       GIVEN
176       PROOF OF: (Perceives! capt t3 (not (Reachable capt t3 lga13)))
177       GIVEN
178       PROOF OF: (R capt t3 (Land capt t3 hud) (Land capt t3 teb1))
179       Applied 'Reasonableness Definition' to: (Perceives! capt t3 (Reachable capt t3 hud)); (Perceives! capt t3 (not (Reachable capt t3 teb1)))
180         PROOF OF: (Perceives! capt t3 (not (Reachable capt t3 teb1)))
181         GIVEN
182         PROOF OF: (Believes!1 capt t3 (Land capt t3 teb1))
183         Applied 'Level-1 Definition' to: (Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
184           PROOF OF: (Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
185           Proved via ShadowProver:
186           -->(AssumptionsNowContainsGoal

```

```

187 --> Givens:
188 -->(((Perceives! capt t3 (Reachable capt t3 hud))
189 --> (GIVEN[])))
190 -->
191 -->((Knows! capt t3 (Reachable capt t3 hud))
192 --> (Perception to knowledge (Perceives! capt t3 (Reachable capt t3 hud)))
193 --> [(GIVEN[])])
194 -->
195 -->((Believes! capt t3 (Reachable capt t3 hud))
196 --> InferenceJustification{base=[(Knows! capt t3 (Reachable capt t3 hud))], message='DR5'})
197 -->
198 -->((Perceives! capt t3 (not (Reachable capt t3 teb1)))
199 --> (GIVEN[])))
200 -->
201 -->((Knows! capt t3 (not (Reachable capt t3 teb1)))
202 --> (Perception to knowledge (Perceives! capt t3 (not (Reachable capt t3 teb1))))
203 --> [(GIVEN[])])
204 -->
205 -->((Believes! capt t3 (not (Reachable capt t3 teb1)))
206 --> InferenceJustification{base=[(Knows! capt t3 (not (Reachable capt t3 teb1))], message='DR5'})
207 -->
208 -->((Perceives! capt t3 emergency)
209 --> (GIVEN[])))
210 -->
211 -->((Knows! capt t3 emergency)
212 --> (Perception to knowledge (Perceives! capt t3 emergency)
213 --> [(GIVEN[])])
214 -->
215 -->((Believes! capt t3 emergency)
216 --> InferenceJustification{base=[(Knows! capt t3 emergency)], message='DR5'})
217 -->
218 -->((Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1)))
219 --> (GIVEN[])))
220 -->
221 -->((Knows! capt t3 (Says! atc t3 (Land capt t3 teb1)))
222 --> (Perception to knowledge (Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1))))
223 --> [(GIVEN[])])
224 -->
225 -->((Believes! capt t3 (Says! atc t3 (Land capt t3 teb1)))
226 --> InferenceJustification{base=[(Knows! capt t3 (Says! atc t3 (Land capt t3 teb1))), message='DR5'])
227 -->((Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
228 -->
229 -->((R capt t3 (Land capt t3 hud) (Land capt t3 lga13))
230 --> (GIVEN[])))
231 -->
232 -->((Perceives! capt t3 (not (Reachable capt t3 lga13)))
233 --> (GIVEN[])))
234 -->
235 -->((Knows! capt t3 (not (Reachable capt t3 lga13)))
236 --> (Perception to knowledge (Perceives! capt t3 (not (Reachable capt t3 lga13))))
237 --> [(GIVEN[])])
238 -->
239 -->((Believes! capt t3 (not (Reachable capt t3 lga13)))
240 --> InferenceJustification{base=[(Knows! capt t3 (not (Reachable capt t3 lga13))], message='DR5'})
241 -->
242 -->((R capt t3 (Land capt t3 hud) (Land capt t3 teb1))
243 --> (GIVEN[])))
244 --> Goals:
245 -->((Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1))))
246 PROOF OF: (Perceives! capt t3 emergency)
247 GIVEN
248 PROOF OF: (Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1)))
```

```

219    -->  (GIVEN[]))
220    -->
221    -->((Knows! capt t3 (Says! atc t3 (Land capt t3 teb1)))
222      (Perception to knowledge (Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1))))
223      [(GIVEN[])])
224    -->
225    -->((Believes! capt t3 (Says! atc t3 (Land capt t3 teb1)))
226      InferenceJustification{base=[(Knows! capt t3 (Says! atc t3 (Land capt t3 teb1))), message='DR5']}
227      (Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
228    -->
229    -->((R capt t3 (Land capt t3 hud) (Land capt t3 lga13))
230      (GIVEN[]))
231    -->
232    -->((Perceives! capt t3 (not (Reachable capt t3 lga13)))
233      (GIVEN[]))
234    -->
235    -->((Knows! capt t3 (not (Reachable capt t3 lga13)))
236      (Perception to knowledge (Perceives! capt t3 (not (Reachable capt t3 lga13))))
237      [(GIVEN[])])
238    -->
239    -->((Believes! capt t3 (not (Reachable capt t3 lga13)))
240      InferenceJustification{base=[(Knows! capt t3 (not (Reachable capt t3 lga13))), message='DR5']}
241    -->
242    -->((R capt t3 (Land capt t3 hud) (Land capt t3 teb1))
243      (GIVEN[])))
244    --> Goals:
245    -->(Believes! capt t3 (Believes! atc t3 (Land capt t3 teb1)))
246    PROOF OF: (Perceives! capt t3 emergency)
247    GIVEN
248    PROOF OF: (Perceives! capt t3 (Says! atc t3 (Land capt t3 teb1)))
249    GIVEN
250
251 Calling Adjudicator...
252 Final Belief: (Believes!2 adj t3 (Land capt t3 hud))
253 Hence: (Knows! adj t3 (Ought! adj t3 (happens (action adj land(hud)) t3' emergency)))
254
255 Generating a solution to Sub-Argument 4...
256 Proof found in 0.9096662998199463 seconds.
257 PROOF OF: (happens (action capt (land hud)) t4)
258 (AssumptionsNowContainsGoal
259 Givens:
260 (((Knows! capt t3 (Ought! capt t3 emergency (happens (action capt (land hud)) t4)))
261   (GIVEN[]))
262 (Ought! capt t3 emergency (happens (action capt (land hud)) t4))
263
264 ((Believes! capt t3 (Ought! capt t3 emergency (happens (action capt (land hud)) t4)))
265   InferenceJustification{base=[(Knows! capt t3 (Ought! capt t3 emergency (happens (action capt (land hud)) t4))), message='DR5']}
266 (happens (action capt (land hud)) t4)
267 (Intends! capt t3 (happens (action capt (land hud)) t4))
268
269 ((Perceives! capt t3 emergency)
270   (GIVEN[]))
271
272 ((Knows! capt t3 emergency)
273   (Perception to knowledge (Perceives! capt t3 emergency)
274   [(GIVEN[])])
275
276 ((Believes! capt t3 emergency)
277   InferenceJustification{base=[(Knows! capt t3 emergency)], message='DR5'})
278 Goals:
279 (happens (action capt (land hud)) t4))
280

```