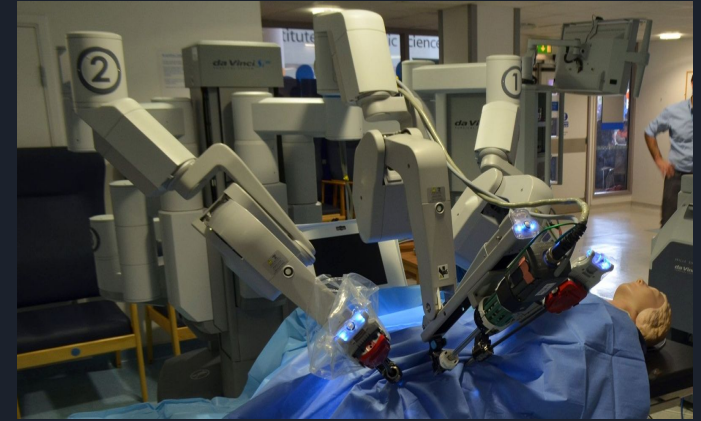


6. The Divine-Command Approach to Robot Ethics

Sergio González Muriel
Antonio David Ponce Martínez

Introducción

- Problema:
 - Cuando una vida humana depende de un robot, ¿Qué debería hacer?
 - ¿Quién tiene la responsabilidad de tal decisión?
- ¿Realidad o ficción?
 - Foster-Miller TALON
 - ATLAS
 - Avances en IA



Códigos éticos

- Reglas de compromiso
 - Acordadas por alguna nación o grupo
- Utilitario
 - Representado en lógica deóntica
 - Basado en lógica computacional
- Éticas de comando divino
 - Basado en los conceptos monoteístas de qué es correcto e incorrecto





Ramas de la ética

- Metaética
 - Estado ontológico de los conceptos básicos de la ética (Correcto e incorrecto)
- Éticas aplicadas
 - Más práctico y específico
 - Comienza con un conjunto de reglas morales para aplicarlas a escenarios concretos
- Éticas normativas (Ética moral)
 - Compara y contrasta formas de definir conceptos de obligación, prohibición, deber, etc.
 - Consecuencialistas
 - No consecuencialistas



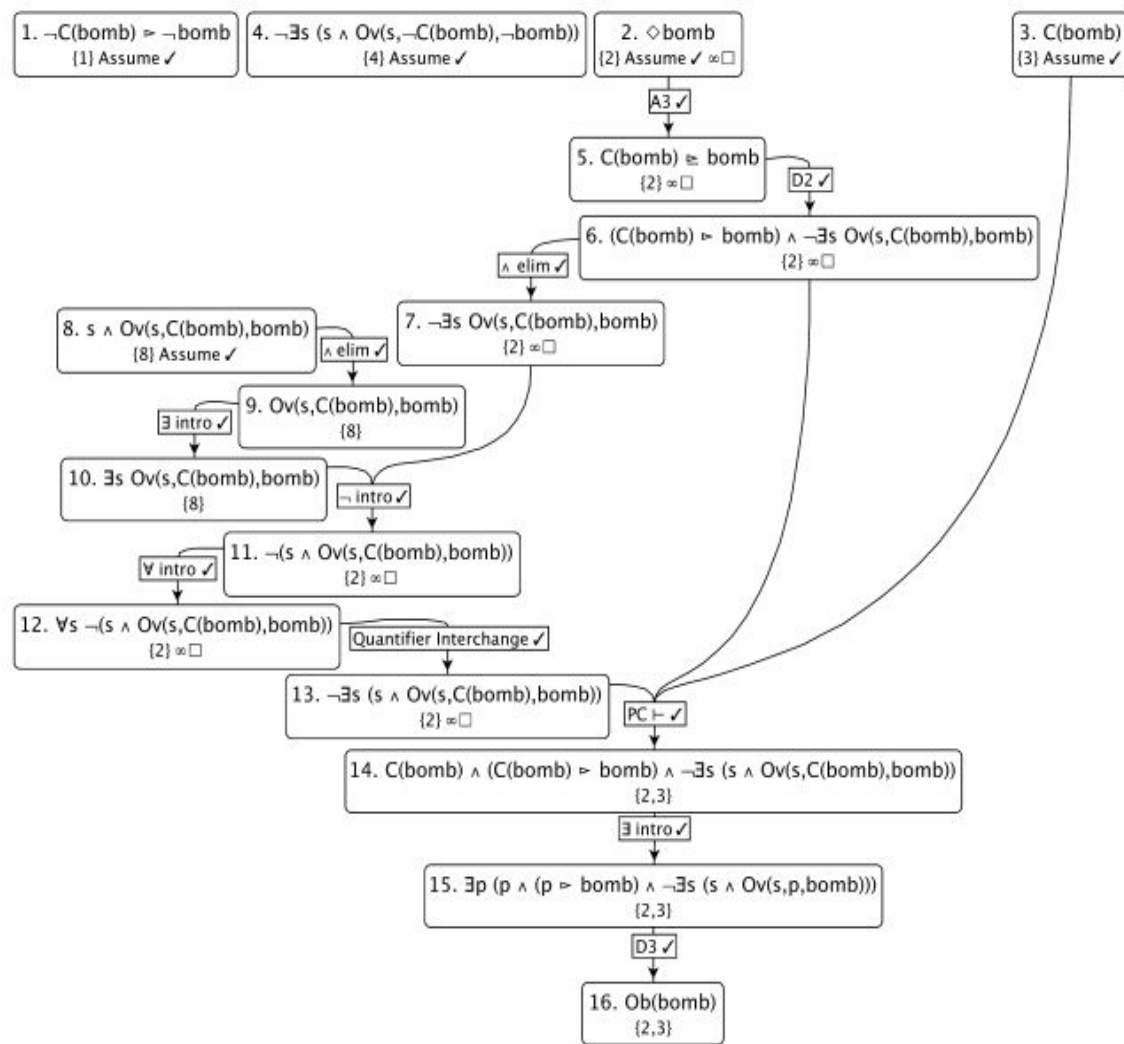
Robots éticamente correctos

1. Sólo realizan acciones permitidas
2. Todas las acciones relevantes que son obligatorias son llevadas a cabo, teniendo en cuenta dependencias y conflictos con otras acciones.
3. Todas las acciones permitidas pueden ser verificadas por el robot como tal, y dicha verificación debe de poder ser expresada en lenguaje natural.



Axiomas, teoremas y definiciones

- A1: $\forall p \forall q pRq \supset M(p \& q)$
- Teorema 1: $\forall p \forall q pRq \supset Mp$
- Teorema 2: $\forall p \forall q pRq \supset Mq$
- A2: $(pRq \& pRs) \supset pR(q \& s)$
- D1: $sOpq = \text{def } pRq \& \sim((p \& s)Rq) \& M(p \& s \& q)$
- D2: $plq = \text{def } pRq \& \sim \exists s(sOpq)$
- D3: $Oq = \text{def } \exists p(p \& pRq \& \sim \exists s(s \& sOpq))$
- A3: $Mp \supset (Cp)Ip$



- **A1:** $\forall p \forall q pRq \supset M(p \& q)$
- **Teorema 1:** $\forall p \forall q pRq \supset Mp$
- **Teorema 2:** $\forall p \forall q pRq \supset Mq$
- **A2:** $(pRq \& pRs) \supset pR(q \& s)$
- **D1:** $sOpq = \text{def } pRq \& \sim((p \& s) Rq) \& M(p \& s \& q)$
- **D2:** $p \supset q = \text{def } pRq \& \sim \exists s (sOpq)$
- **D3:** $Oq = \text{def } \exists p (p \& pRq \& \sim \exists s (s \& sOpq))$
- **A3:** $Mp \supset (Cp) Ip$