Optimalitätstheorie ausserhalb der Phonologie

Jochen Trommer

jtrommer@uni-leipzig.de

Universität Leipzig
Institut für Linguistik

Einführung in die Phonologie – WS 2006/2007

Ethics for Robots: An illustration of OT

Isaac Asimov described what became the most famous view of ethical rules for robot behaviour in his "three laws of robotics:"

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law.

Ethics for Robots: An illustration of OT

Isaac Asimov described what became the most famous view of ethical rules for robot behaviour in his "three laws of robotics:"

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings, **except where such orders would conflict with the First Law**.
3. A robot must protect its own existence, **as long as such protection does not conflict with the First or Second Law**.

Ethics for Robots in OT

*INJURE HUMAN : A robot may not injure a human being or, through inaction, allow a human being to come to harm.

OBEY ORDER: A robot must obey the orders of human beings.

PROTECT EXISTENCE: A robot must protect its own existence.

Ranked:

*INJURE HUMAN: ≫ OBEY ORDER ≫ PROTECT EXISTENCE

Story A:

Human says to **R**obot: Kill my wife!

1. **R** kills **H**'s wife
2. **R** kills **H** (who gave him the order)
3. **R** doesn't kill anyone
4. **R** kills himself Second Law.

	*INJURE HUMAN	OBEY ORDER	PROTECT EXISTENCE
R kills **H**'s wife	*!		
R kills **H**	*!	*	
☞ **R** doesn't kill anyone		*	
R kills himself		*	*!

Story B:

Human says to **R**obot: Kill my wife or I kill her!

	*INJURE HUMAN	OBEY ORDER	PROTECT EXISTENCE
☞ **R** kills **H**'s wife	*		
R kills **H**	*	*!	
R doesn't kill anyone	*	*!	
R kills himself	*	*!	*

Story C:

Human says to **R**obot: Kill my wife or I destroy you!

	*INJURE HUMAN	OBEY ORDER	PROTECT EXISTENCE
R kills **H**'s wife	*!		
R kills **H**	*!	*	
☞ **R** doesn't kill anyone		*	*
☞ **R** kills himself		*	*