

ROBOTS CAN BE MADE WITH HUMAN-LIKE  
EMOTIONS—BUT SHOULDN'T BE

John McCarthy

Computer Science Department

Stanford University

jmc@cs.stanford.edu

<http://www-formal.stanford.edu/jmc/>

March 22, 2004

A slogan for AI:

Whatever a person can do, the person should be able  
make a computer do for him.

Almost all of my papers are on the web page.

## USEFUL ROBOTS WON'T HAVE HUMAN-LIKE EMOTIONS

- human motivations and emotion
  - direct reaction to inputs
  - chemical state
  - subgoals escaping from main goal
- Robot and Baby story  
<http://www.formal.stanford.edu/jmc/robotandbaby.h>
- hatred. Police robots shouldn't hate criminals.
- depression, joy - I can't do them now, but maybe can.
- robots in society. Robots shouldn't be capable of suffering or personal ambition.
- Computer understanding emotion requires logical formalization.

## ONE PERSON HATES ANOTHER

- If *person1* hates *person2*, *person1* prefers situation which *person2* is worse off.
- $Hates(person1, person2, s)$
- $Hates(person1, person2, s) \wedge WorseOff(person2, Result(e, s), s) \rightarrow Prefers(person1, Result(e, s), s)$