Don't Think of a Killer Robot: Framing the Public Perception of AI

(Remarks for IJCAI 2015 panel on "Who Speaks for AI")

Subbarao Kambhampati

7/29/2015

Note: These were just speaker notes; the panel itself consisted of oral statements (audio version included next to this link)
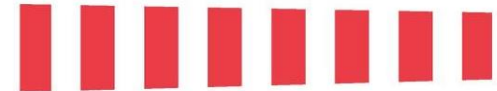
Don't Think of a Killer Robot..

- As I do my 5 min spiel here, I have but request to you all. Don't think of a killer robot. Whatever you do, don't think of a killerbot!
  - If you actually managed to not think of killer robots, you must be an AI agent.. Let me explain..
- George Lakoff writes that he starts his linguistics class every year by asking his whole class "Don't think of an elephant; whatever you do, don't think of an elephant"
  - And proceeds to ask the students what they are thinking of

DON'T THINK OF AN ELEPHANT!

NEW YORK TIMES BESTSELLER

KNOW YOUR VALUES AND FRAME THE DEBATE

GEORGE LAKOFF

So, what has this got to do with the panel?

Framing the Debate..

- The good news is that we are finally dominating the news waves;
- The worrisome news is that much of the coverage is being framed in terms of how to stop the Evil AI rather than about the many benefits of AI.
  - I realize some of this is certainly because of the overzealous reporting—but we could also do a better job.
  - ..and lord knows we have been a bit starved for public attention since the heyday of expert systems
    - I mean the whiplash inducing jerk with which popular opinion went from "AI is a dud" to "AI is going to take over" is something to behold
- But this is certainly starting to have some unexpected consequences. I have heard an federal program manager tell me that they are being asked to be careful about mention AI in senate testimonies etc.
- I think we need to take better lead in framing the public discourse on AI.
  - We are currently following rather than taking the lead..
  - We should resist

Theme of IJCAI 2016
Human-Aware AI

- It is important that we focus on emphasizing the societal benefits of AI before dwelling on the possible safety/security issues
  - There is a of course a lot we can do, and are doing—AAAI openhouse for example
- This is why, for IJCAI 2016 next year, we are going to have "Human-Aware AI" as the special theme
  - While I probably could have gone with a "Stopping Evil AI" special track, and raised a lot of sponsorship and attention, we decided to chose this more general theme
  - chose this specifically to highlight the many positive interactions AI systems will continue to have in cohabiting with humans
    - Including human-robot teaming,
    - human-in-the-loop decision support systems,
    - Human-machine collaboration (embodied or virtual)
    - It also pushes the frotiers of AI research in many interesting ways—including those who have drifted to hjman factors etc

On Defragmenting AI

- In IJCAI 2016, we will be going back to a single main-track

  - No <u>horizontal special tracks</u> such as "Planning" track, "Learning" track, "KR" track etc.

- Rationale: If we steadfastly stick to our own silos even at IJCAI, when will get to focus on integrated AI?