



Validation and Verification of Artificial Intelligence Systems

Where do We Stand?

Ahmed Kamel
kamel@cord.edu



CONCORDIA COLLEGE
MOORHEAD, MINNESOTA

Brief Bio



Dr. Ahmed Kamel is a Professor of Computer Science and Management Information Systems at Concordia College in Moorhead, Minnesota, USA. He received a Ph.D. in Computer Science from Michigan State University in 1994 specializing in Artificial Intelligence. He also received an MBA from North Dakota State University in 2013. Dr. Kamel's research interests in Applied Artificial Intelligence and Knowledge-Based Systems are broad. Past research included diverse application areas including representation of and reasoning about physical devices, software support for engineering design, support for planning for agricultural farming, integration of linked problem solving agents, and office automation. His interests also include social impact of computers, and their applications. Dr. Kamel has over sixty publications on various aspects and applications of Artificial Intelligence.

Compact History of AI

- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1950s Early AI programs, including Samuel's checkers program,
- Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1956 Dartmouth meeting: "Artificial Intelligence" adopted
- 1965 Robinson's complete algorithm for logical reasoning
- 1966-74 AI discovers computational complexity
- Neural network research almost disappears

Compact History of AI (Cont.)

- 1969-79 Early development of knowledge-based systems
- 1980-88 Expert systems industry booms
- 1988-93 Expert systems industry busts: “AI Winter”
- 1985-95 Neural networks return to popularity
- 1988- Resurgence of probability; general increase in technical depth, “Nouvelle AI”: ALife, GAs, soft computing
- 1995- Agents, agents, everywhere : resurgence of expert systems
- 2003- Human-level AI back on the agenda

Artificial Intelligence

Unfortunately, the biggest disservice ever to the field of Artificial Intelligence was naming it with that name: it raises unreasonable expectations.

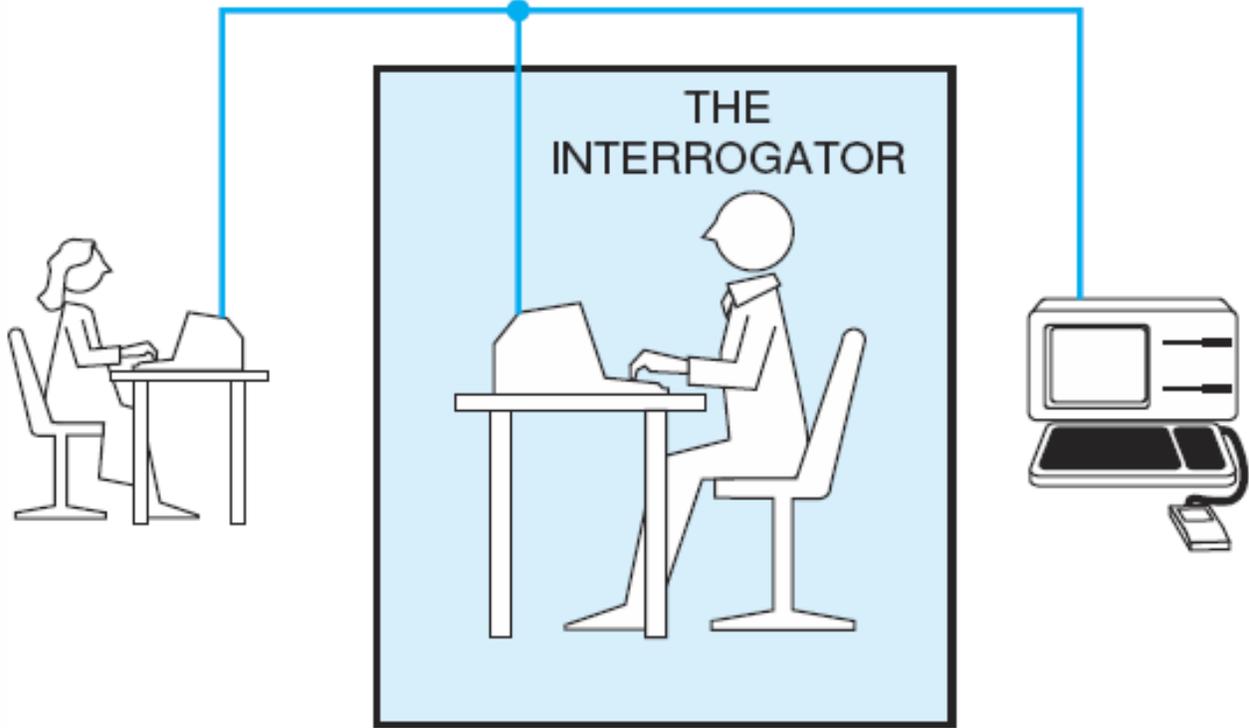
Artificial Intelligence is not about making Computers Intelligent but about making them **seem** intelligent which is at best a big hurdle for V&V

The Turing Test

Alan Turing devised a test to be used as the basis for judging whether or not a Computer System is Intelligent:

- Have a person interrogate a computer and another person only through a keyboard without knowing which is which. If the person cannot distinguish the computer from the other person, the computer passes the test

The Turing Test



The Turing Test (cont.)

Problem: Developers started writing programs that exhibit human behavior (like being slow to respond), thus defying the purpose

Do we really need this level of *sophistication* in computer programs?

What does this mean for V&V?

Future of AI

Are we going to see Machines that are smarter than humans?

- There is no easy answer
- If this is the case what does V&V look like? Whow would it be performed by?

People used to say that it would be impossible to build machines that are physically stronger than humans. Clearly, this has long been a reality.

The Physical Symbol System Hypothesis

- Many AI researchers contend that a physical symbol system (PSS) is necessary to represent intelligence.
- A PSS is defined as:
 - A set of symbols representing objects from the real world
 - A set of rules for manipulating those symbols
- This excludes newer approaches like neural networks

Approaches to AI

- 2 camps and the continuum in the middle:
 - engineering point of view: method is not important as long as it achieves the desired result
 - cognitive point of view: method is the primary concern
- Each of these approaches would have very different implications for V&V

Some Areas of AI

- Game playing; some games such as chess have always been associated with intelligence
 - Since game playing is inherently competitive without a well-defined end goal, what does V&V even mean?

Some Areas of AI (cont.)

- Expert Systems: defined as computer systems that mimic the knowledge of a human expert (in some domain).
- To date, this is the most successful area of AI accounting for most applications. Applications span a wide range of areas (medical diagnosis, mechanical diagnosis, farm management, military planning, financial forecasting and planning, engineering design, ...etc.)
- This is the area with the highest expectation for V&V. However, problem solving is inherently heuristic which at best limits V&V
 - Most developers don't even attempt V&V beyond observational comments about the quality of resulting recommendations

Some Areas of AI (cont.)

- Intelligent Agents:
 - One of the newest areas of symbolic of AI
 - A collection of collaborating agents
 - Each agent is autonomous and situated
 - Essentially expert systems that are autonomous and collaborative
 - Suffer from same V&V limitations like expert systems

Some Areas of AI (cont.)

- Natural Language Understanding: Wouldn't it be nice if we can talk to Computers in English rather than in cryptic languages?
 - It is impossible to even define V&V:
 - Natural language is inherently vague

Some Areas of AI (cont.)

- Machine Learning: Wouldn't it be nice if Computers would learn from their own mistakes?
- Machine learning falls into 2 categories:
 - Symbolic machine learning relying on feedback to add more rules/logic
 - Addition to expert systems with same V&V limitations
 - Connectionist approaches (Neural Networks): diverting from physical symbol systems and attempting to model the interactions inside our brains
 - Relies on our limited understanding of human neural functions
 - Hard to validate something we do not fully understand
 - Further complicated by black-box approach both to problem solving and to learning

Conclusions

- Validation and Verification in Artificial Intelligence lags behind other software disciplines
- V&V within symbolic AI is limited to heuristically validating execution results
- V&V in connectionist AI is virtually non-existent
- A big question remains: Given the nature of AI, is V&V even desirable?