VIEW POINT

EMBRACING AI: MANAGING THE UNKNOWN

Infosys®
Navigate your next

Can we imagine matching fingerprints in forensic analysis without machines? How would a logistic company optimize its cost and still deliver a million packages a second across the globe? Without a doubt, Artificial Intelligence (AI) and Machine Learning (ML) bring value to our lives. Machines are simply better at solving certain problems and whether we like it or not, the reality is that machines are getting smarter. Per Elon Musk, failing to acknowledge this fact i.e. machines are getting smarter is what makes humans vulnerable.[1]

Mankind can be a beneficiary of AI and ML only if it ”befriends” those machines that improve human life. Conflict possibly starts when machines begin to learn and cognize in a way that we cannot understand. In other words, conflict begins when machines begin to think. But can machines think? Alan Turing explored this very question in his path breaking paper published in 1950 concluding that such a possibility does exists[2]. How does such a system, a “thinking” machine, pose a problem? Consider the Lion Air crash[3] that happened in 2018 and how the pilots struggled to override the Maneuvering Characteristics Augmentation System (MCAS). One of the reasons they struggled is because they were not aware of how the system functioned. While the machine, the aircraft in this case may not have been thinking, the point I want to focus is the perspective of the crew. Many precious lives were lost. Definite risk arises when dealing with a system (AI or non-AI) whose behavior we cannot understand.

External Document © 2020 Infosys Limited

As machines begin to 'think' in a way that is different from how humans think, how should we evolve such a system? What should a machine be allowed to do and learn? Most importantly, how should any potential risk be mitigated? Solutions we seek through technology have 3 important characteristics that are listed below. (see Figure 1).

1. being intelligible to humans

2. something that machines can provide

3. involving cognition

Some solutions involve cognition, some solutions are intelligible to humans, some solutions can be provided by machines; However, the most interesting solutions emerge when these characteristics intersect.

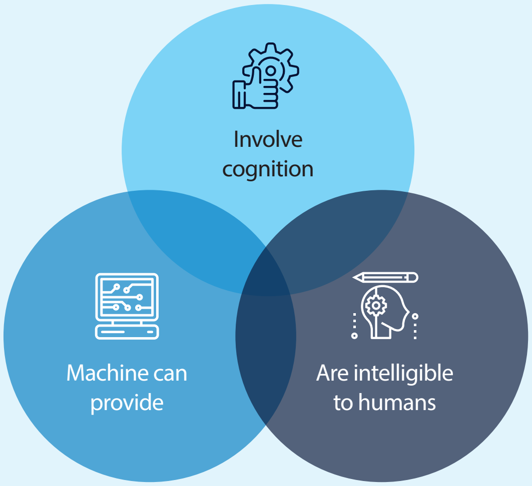

Figure 1 Solution Characteristics

Robots (automate):

The intersection of solutions that are intelligible to humans and those that machines can provide, shown in Figure 2, is possibly the simplest scenario. A useful outcome that does not involve cognition and that machines can provide is something with which robots can help us. Robots relieve humans of manual labor, allowing us to pursue other, more complicated tasks. We can characterize robots in the following ways:

- They lack cognition
- They automate, i.e., they accomplish what humans accomplish through manual actions
- How they behave is intelligible to humans

Robots need to be deployed in ways that are safe for humans. In summary, robots can be leveraged to automate tasks that do not involve cognition.

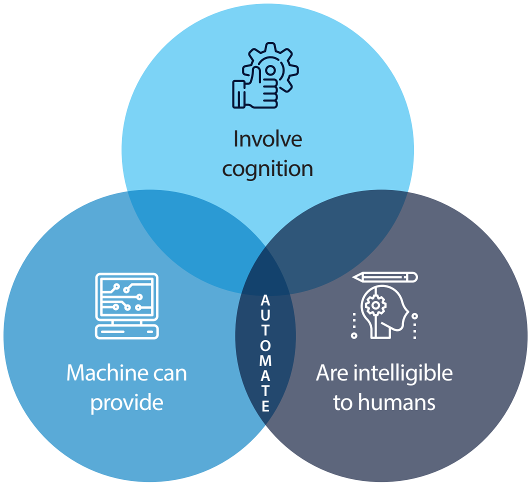

Figure 2

External Document © 2020 Infosys Limited

Autonomous agents (delegate):

Figure 3 shows the intersection of solutions that involve cognition and those that machines can provide. How a human accomplishes a task, i.e., his/her technique, is likely to remain opaque to other humans. When we hire someone for a task, we don't try to understand that person's technique. We'd rather define the role clearly and let that person focus on how to achieve the outcome. Legal and ethical boundaries, in the form of policies, may provide constraints. Can an autonomous agent, a machine, be deployed in a similar manner? What if we have a way to authorize and monitor that agent's access to appropriate resources and have a way to regulate where and how the agent is used?

This intersection space holds all such autonomous agents that involve cognition, whose behavior we may not understand, but whose outcome we will be able to validate. Such agents will have the following characteristics:

- The outcome of the agent is clearly defined.
- How the agent reaches its conclusions need not be intelligible to humans.
- Resources these agents will have access to, will be controlled through policies to enforce legal and ethical boundaries.
- The agent can be turned off at will.

Regulation will play an important role in defining how and where these agents get deployed. For instance, consider an agent trained to detect and recognize human faces. While its usage might be justified from a 'security' standpoint, it may not be so from a 'tracking' standpoint. Autonomous agents are to be deployed to carry out tasks that involve cognition and whose outcomes humans can override, if needed.

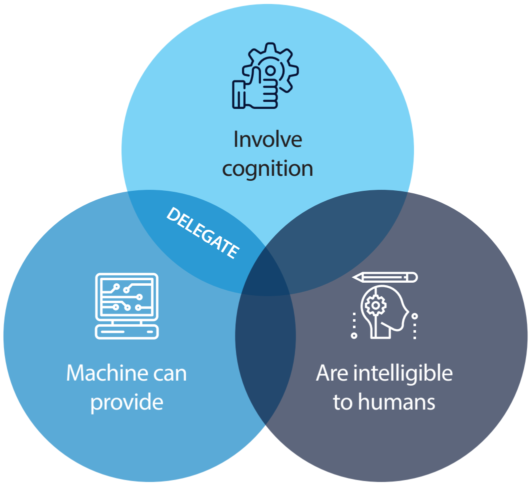

Figure 3

External Document © 2020 Infosys Limited

The lieutenants (decide):

The intersection of solutions shown in Figure 5 are intelligible to humans, involve cognition, and can be provided by machines. Machines that offer such solutions can be deputies or decision assistants—lieutenants and can help humans come to decisions. The crucial aspect here is clarity around how a decision is reached. In other words, the decision needs to be traceable and explainable to the human who ultimately owns the decision. These systems need to be wired (with or without AI) in such a way that they can trace out the decision path concerning a complex scenario as a graph or a decision tree that can be consumed by the decision maker. The nodes in the graph can be autonomous agents but the edges need to be visible. I consider the Decision Review System (DRS)[4] deployed in the game of cricket a classic example of a lieutenant in action. The scenario of Leg Before Wicket (LBW)[5] is a very subjective decision and a complex one to automate with many influencing factors. While the decision-making process in the DRS might involve an autonomous agent to predict the trajectory of the ball, whose technique may not be fully explainable, the decision process overall is very much traceable and consumable through a visual. In summary, these are the characteristics of lieutenants:

- They help with decision making that involves complex scenarios.
- How they arrive at a decision must be intelligible to humans.
- Their behavior can be expressed as a graph or a tree.
- The nodes may or may not include autonomous agents.

Lieutenants are to be deployed to help humans with decision making.

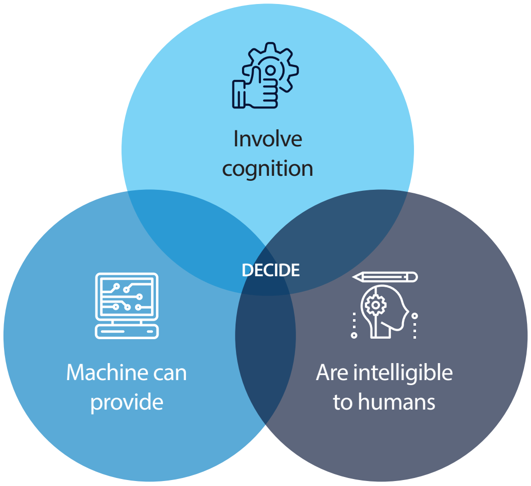

Figure 5

External Document © 2020 Infosys Limited

External Document © 2020 Infosys Limited

The navigators (navigate):

The intersection of solutions shown in Figure 6 includes outcomes that involve cognition in a way that is intelligible to humans and which machines cannot provide. Either machines are not yet trained to produce these solutions, or humans have not yet found a way to program the machines. Outcomes that result from imagination, intuition, and gut feeling fall into this space. What could these outcomes address?

These outcomes could help us answer the following questions:

- How can we measure and evolve autonomous agents?
- How best to design our lieutenants?
- How should we design our policies to constrain our autonomous agents into behaving ethically and morally?
- How do we recognize and eliminate biases in AI and ML systems?

In summary, our intuition and imagination will help us navigate into the future.

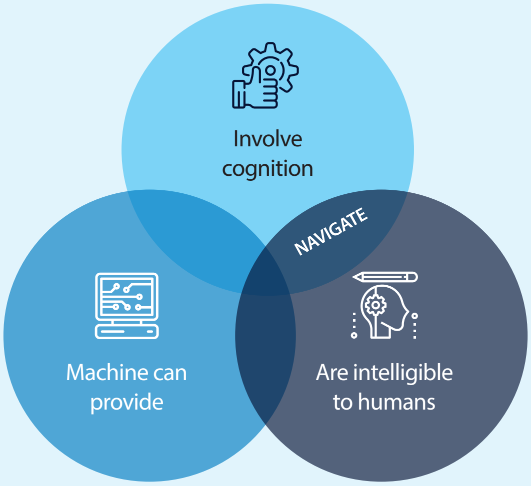

Figure 6

Conclusion:

In the context of our original set of questions - If machines 'think' in a way that is different from how we think, how should humans evolve such systems? What should a machine be allowed to do and learn? Most importantly, how should risk be mitigated?

1. We should use machines i.e. robots to automate (with care) what does not involve cognition, within the constraints of what machines are capable of accomplishing.

2. We could delegate (and regulate) tasks that involve cognition and whose outcomes are clearly defined to autonomous agents. Such agents should be allowed to learn on their own as long as we have the capacity to override agents' outcomes and to control the resources to which they have access.

3. We could use lieutenants that are designed to explain how they arrive at a decision to help us with decision making.

4. We must use our intuition and imagination to measure the efficiency of the agents, evolve the AI ecosystem, and navigate into the future.

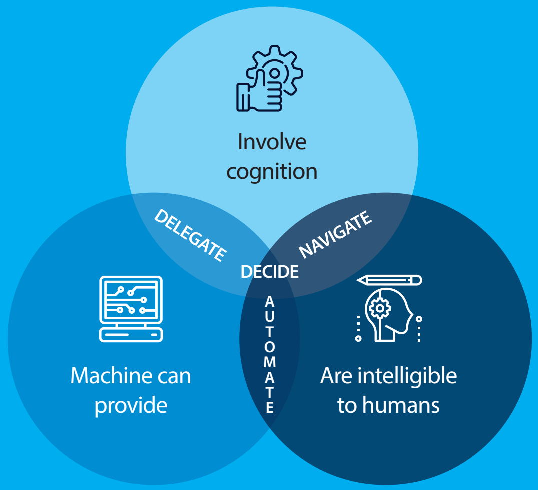

Figure 7 Solution characteristics & usage contexts

External Document © 2020 Infosys Limited

About the Author

Ramaswami Mohandoss
Senior Architect, DNA, Infosys

Ramaswami Mohandoss is a seasoned architect and a data enthusiast. He works at the intersection of architecture and data science and has rich Analytics, Big data and Cloud experience. In the past decade, he has worked with several Retail clients and has helped them monetize their data assets through effective technology solutions and data products.

References

1. https://www.youtube.com/watch?v=yG6Ay3nlB84
2. https://www.csee.umbc.edu/courses/471/papers/turing.pdf
3. https://en.m.wikipedia.org/wiki/Lion_Air_Flight_610#Preliminary_report
4. https://www.hawkeyeinnovations.com/sports/cricket
5. https://en.wikipedia.org/wiki/Leg_before_wicket

For more information, contact askus@infosys.com

Infosys® Navigate your next

© 2020 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.