Commonplace

Ethical Considerations: Applying Natural Systems to AI for Better Situational Understanding

Sharena Rice

Published on: Mar 23, 2021

DOI: 10.21428/6ffd8432.d47f1790

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

“It’s hard for a computer to know what a cup is,” said my lab mate.

His friends laughed at the statement— how could something so trivial to us humans be so hard for computers to understand? A cup is so simple that humans can accurately identify them from infancy. Most adults hardly have to think to process what is and what is not a cup. Technology is coming closer and closer to being able to take over many cognitively-demanding tasks. If a cup is hard for a computer to classify, then what about more complex classifications, such as recognizing emotional mismatches between a person’s tone of voice and their body language? Why do computers lack this kind of intuition?

In order to have an ethical artificial intelligence, we need to consider our human experience in an ever-increasing technological world in addition to the larger contexts beyond individual or community experiences. Using the human nervous system as an infrastructure and lens into human behavior, we can hopefully learn how to create better and more empathetic technology that better suits our needs.

Throughout our lives, we observe the behaviors of others. We make sense of cause and effect relations, putting things into context in a holistic way that creates rich tapestries of stories, rather than simple vectors of numbers. We tell stories about what we have experienced, we are constantly creating narratives and explanations. **The holistic approach gives humans an advantage over many artificial intelligence systems: they connect cause and effect relationships.**

Humans can use the context beyond a relatively short list of properties to determine why something happened as it did. They can see some things that are invisible to an artificial intelligence because their brains synthesize the information in ways that lead to greater insights than the sum of their parts. While machines may need numerous examples to learn what a cup is, a human may only need one example. A single traumatic experience or an extremely positive one-time experience can change many things about how humans interpret many their situations in life going forward. A major issue that top engineers, data scientists, and software developers need to work with is how to create artificial intelligences that are able to interpret situations similarly to or better than how humans would.

A note on humanity, empathy, and perspective

What makes a human a human? There are three major aspects that I came up with that are universal across living humans. All three must be fulfilled in order to be human, yet all three are fuzzy concepts without absolute definitions:

1. Being alive.
2. Having their own perspective that comes from the inside.
3. Being in the form of a human.

Humans tend to be human-centric because their nervous systems tend to be wired to place greater value on human interests. Humans may also be impulsive in valuing short-term results over long-term outcomes. Thus, I'd argue that even if the human nervous system's capabilities can create better artificial intelligence, "humanizing" artificial intelligence is not the goal. The goal is to create the truest situational understandings for the benefit of ALL beings, not just in the short-term, but also in the long-term. This may be the ideal in a world that contains many more beings than humans. The world also has animals, plants, water, and a larger context. External environmental situations affect reflective inner states. Autonomous vehicles are not truly separate from the places they are located or the people who interact with them. The universe is ultimately an interconnected system. This is part of why **a wide diversity of human and non-human perspectives must be considered in the creation of artificial intelligence. Fortunately, we may be able to use the intelligent capacities of other natural systems to create better artificial intelligences.**

Empathy, the ability to take the perspective of another person, is frequently raised as a part of what makes humans humans. This notion is flawed, as it dehumanizes many kinds of humans. Very young children typically have a harder time taking the perspectives of other people than typical adults do. Does this make very young children any less human than adults? Along similar lines, different adults have different abilities to take the perspective of others. Sometimes, peoples' judgments of other peoples' perspectives are incorrect assumptions due to differences in emotional expression as seen from the outside. Are empaths actually more human than typical people, or do they just seem to be so from the outside because they are using different strategies in their interactions than typical people? And what if it is easier for a highly-sensitive person to empathize with others who are also highly-sensitive because they had a unique shared experience which people with typical sensory systems would not truly understand?

A person could be empathetic towards others without being compassionate: they could take other peoples' perspectives without actually caring about other peoples' well-being. A person could also be compassionate or sympathetic without being empathetic: treating people with great care using an intellectual rather than intuitive understanding of other peoples' likes and dislikes despite having hardship with placing themselves in another person's shoes. At the end of the day, humans cannot truly know what it is like to be someone else unless they have actually been someone besides themselves.

Modern humans live in an increasingly virtual, passive world of disconnect between minds and bodies. From an increased heartbeat while excited, feeling heavy when tired, every emotion naturally has accompanying sensations. Deliberately mirroring another person via the mind and through body language, and taking notice which sensations arise in your own body during this mirroring is a strategy that can allow humans to better interpret the emotions of others. Along with mirroring, attending to the larger situational context beyond another person's present body language and integrating information from both is helpful when interpreting people: doing both is a strategy I call "hypermirroring". Hypermirroring is a challenge to do if inattentive, overwhelmed, fearful, tired, or stressed. But are people in a state of stress any less human than people who are relaxed? Some people experience a sense of separation between mind and body due to living in modern industrialized worlds where bad posture and passive virtual motions tend to be the norm. Are they less human because they have a harder time being fully-present within themselves, and by extension, in their interactions with others?

As soon as some people are considered less "human" than others, this creates artificial divides within and between people via value judgments, resulting in societal ills. In an ideal world, all humans would be respected and treated as humans. When some states within the same individual are considered less "human", this creates an artificial divide where a person disintegrates parts of themselves, resulting in a more troubled mind. In an ideal world, all human states of mind would be acknowledged and treated with care by the individual who experiences these states. **Although empathy is not what makes a human a human, algorithmic proxies for hypermirroring may lead to truer situational analyses and more appropriate responses from artificial intelligence.**

Interestingly, many humans enjoy the company of their pets even though other animals cannot possibly understand the human perspective because they are not in the human

form. Thus, animals are limited in their empathy for humans. Many dogs will be happy to see their human owners regardless of the owners' states of mind, which positively affects the owners. Despite the lack of capacity for true empathy, emotional support animals help humans going through hard times. Artificial intelligences may also benefit people by responding appropriately to situations or by helping people see themselves from another perspective. Given that artificial intelligences are not alive and not in the human form, it would be an enormous challenge to create a useful proxy for guilt upon wrongdoing, a true reflective sense of justice, or a sense of mortality. Although machines are inherently not human, humans may work alongside artificial intelligence as an efficient tool for creating a better world. The artificial intelligence may also help to expand human perspective beyond variables they would realize and store in working memory on their own.

Explainability

Is it ethical to have artificial intelligence that is impossible for humans to explain? Applying our understanding of the nervous system may make for more understandable artificial intelligence. If we start from principles based on our best understanding of the nervous system and principles of nature, then we may be better able to mitigate errors, biases, and other problems that would otherwise be unforeseen in artificial intelligence because it is based on something we can explain. Our understanding of the natural world is imperfect. Fundamental principles of neural circuits and brain rhythms are still under investigation. As the field of neurotechnology advances, we may reach a greater understanding of the truths of how the nervous system works and advance the field of artificial intelligence while maintaining the ability to explain how the artificial intelligence works.

Principles of the nervous system may be applied to create better artificial intelligence systems for situational understanding. They may also allow humans to continue understanding artificial intelligence and decrease risks associated with misunderstanding artificial intelligence.

Why is the nervous system a good natural system to apply to artificial intelligence?

- **It senses signals.** Like some machines, the nervous system is equipped with an variety of sensors that take in information.
- **It processes a filtered subset of these signals and changes according to the situation.** Jumping into a cool pool from a hot jacuzzi or vice versa will trigger

strong signals about the change in temperature, but when the nervous system gets used to the temperature, then it becomes less alarmed. Similarly, you may notice the smell of coffee upon entering a coffee shop, but after being in the coffee shop for a while, you will stop noticing this scent unless you choose to give more attention to it. Similarly, detecting changes in sensory information, rather than constantly computing unchanging sensory information may be applied for signal processing in artificially-intelligent devices.

- **It integrates signals.** Multisensory information can give a more accurate interpretation of the situation than one sense operating alone. Sensory substitution may take place, as when some persons with blindness [“see” their surroundings](#) by using clicks or when people who are hard of hearing [feel sounds](#) through vibration.
- **It performs computations.** This helps to make decisions by weighing options.
- **It gives outputs as responses to stimuli.** Some reflex arcs allow for responses in the peripheral nervous system before being sent to the central nervous system. The knee jerk response to the physician’s mallet happens even if a person tries to resist it.
- **It learns and remembers.** There are feedback loops for everything from postural control to skill acquisition to memorization. The nervous system generates teaching signals, which compare the actual result to the desired result, allowing for learning through practice. A child does not need to study physics in order to know how to ride a bicycle or to play a game of catch, but they do need practice to do these well. **It functions well enough to have helped every species with a nervous system to survive.**
- **It is efficient in energy consumption, timing, and in size.** While the nervous system is very [power-hungry relative to its weight](#) for a biological organ system, the brain was estimated to be [30 times faster than the world’s fastest supercomputers](#) in 2015. Another estimate says that [the human brain is 10 million times slower than a computer](#) based on the number of computations that can be performed at once, but the computer’s computations are done in serial, compared to the brain’s processing both in serial and mostly parallel. This makes the brain “smarter” at many problem solving tasks. The folds and valleys of the brain allow many neurons to reside in a compact form factor in our heads.

While the nervous system has served humanity well, it is not perfect. It makes mistakes that a machine would not make. Minds sometimes go blank at critical moments or succumb to emotional pressures. A **major problem with the human brain is the tendency towards distraction**, especially in a world full of attention-grabbing notifications or in a state of boredom. Fluctuating between understimulation

and overstimulation makes it challenging for a person to truly stay focused. **Humanity also contains a wide diversity of minds:** different people have different interests, talents, and preferences. Just because two different people are dealt the same hand of cards does not mean they would play them similarly. As it is a good thing for people to have jobs and roles that match who they are, it is a good thing for artificial intelligence to be suitable for the problems that it is employed to solve. **States and traits are also important considerations.** As different people are best suited to different scenarios in certain brain states, artificial intelligence leaders must carefully think through what kind of artificial intelligence traits and what kind of state-switching capabilities would best suit artificial general intelligences that perform a variety of different functions.

Human behavior analysis is a major obstacle for the adoption of autonomous and fleet vehicle technology. In an increasingly-distracted world, having machines that understand human behavior is more vital for human safety than ever. When people move from place to place, they do not behave like tumbleweeds. They can make split-second decisions that completely change their trajectories. An unexpected projectile can fall out of the sky and startle them. People trip. People have reflexes. People realize they forgot something and turn back around. People on the influence of drugs can walk around erratically. Some people use bicycles, some use Segways, some use scooters, some use unicycles. People interact with people and their environments, finely tuning their behavior and their movement patterns, often on a subconscious level. Human behavior is complex, as are the possible situations that influence human behavior.

A big overarching strategy here is: consider how nature has already solved problems in navigation and decision-making and how we can apply these principles to create artificial intelligence systems.

As the Co-Founder and Chief Scientific Officer of [Intvo Inc.](#), this has been my approach to technology development. An interdisciplinary approach that encompasses the diversity of human experiences across cultures and regions would make for better artificially-intelligent road safety systems.

In my dissertation work in the [OmarLab](#), I have studied the neural correlates of navigationally-relevant behaviors. Navigation is a complex, yet evolutionarily conserved task. It involves the integration of multisensory information, movement, and the ability to orient in space. Most animals with a healthy brain and the ability to move are able to find their way home. Even the tiny brains of butterflies allow them to keep track of space as they migrate according to the seasons. The brain creates [inner maps](#)

[of the outer world](#) in ways that are remarkably elegant. There are still numerous things that we have yet to understand about its workings. While conducting my experiments, I have reflected upon **what it means to understand something or someone: to link cause and effect in a way that can influence actions. Truly understanding something means knowing it beyond the words and beyond the numbers. Understanding means creating something like intuition.**

Can we create systems that have some semblance of intuition? Systems that go beyond being able to classify objects and predict situations, such that the artificial intelligence system itself is based on principles of natural sciences? This is an idea behind artificial neural networks, which learns through examples and uses layers of connection for output from one neuron to another. But there are other properties of natural systems that may also be applied to artificial intelligence and computer vision systems. Here are some examples:

- The visual cortex becomes **more active** when a mammal moves **at faster speeds**. When analyzing data using computer vision in real time, it may help in data and energy efficiency to adjust the frames per second according to the speed of the stimulus. If we are creating a camera system at a road infrastructure that gives warnings to the driver based on risk of pedestrian collisions in real time, it would be more helpful for the system to increase its frame rate and become more active if it senses a pedestrian in its visual field.
- **Half of the brain of a dolphin can sleep while the other half is awake**, such that the dolphin can constantly swim. Perhaps this can be applied to space telescopes whose solar panels have somehow been blocked from light for too long, such that the telescope keeps moving while always somewhat “on”.
- The immune system is evolutionarily older than the nervous system with an impeccable memory and great complexity that can give resilience to living beings. Thus, **artificial immune systems may be better models to handle some problems than artificial neural networks are**. Natural immune systems come with their own host of problems, such as the potential for autoimmune disease, but knowledge of natural immune systems may allow us to mitigate these problems in artificial immune systems. Artificial immune systems are under development.

If we could create better situational understandings of broad contexts through combining the best, most suitable aspects of the natural world with the optimized power of machines, this may make for a better, safer world. It would allow us to understand the world and ourselves more truly, which perpetuates yet more questions

and more answers both in artificial intelligence engineering and research strategies to expand knowledge and mitigate the risks that come from barking up the wrong tree.

This would be something for all beings to raise well-classified cups to.

This story was originally published on Medium.com, curated in Dev Genius.