

An Alien “Drove” my car

Online Edition

The final version formatted for the magazine on the IEEE site is available [here](#).

Page 10 Bits Versus Electrons

An Alien Drove My Car: A Cabbie from Mars?

By Bob Frankston

In September 2016, I rented a Tesla with autosteering for the second time. To a degree, I could think of the car as a taxi driving me around, but a taxi driven by an alien intelligence who only acted like a cabbie.

The previous time I rented a Tesla with autosteering was in California, with its well-marked lanes and wide highways. I actually did the driving, but I was getting directions from the navigation system, and much of the steering was done by the car itself, so it's useful for me to think ahead to a time when the car is driving itself rather than just assisting me. We are on a trajectory to fully self-driving vehicles. Driving from Boston, Massachusetts, to New York was very different. The lane markings are designed for people and not easily grokked by the car's steering system.

The car isn't driving itself. It can stay in a lane but relies on me to follow the actual route. It's a symbiotic relationship. At one point, after the car (OK, me, the driver) was led astray by the navigation directions, I found myself on I-384 and needed to get back on I-84 to get to Boston, but the navigation system told me to get on I-291 because it saw I-291 on the sign. A few feet on the same path to I-291/I-84, it told me to get on I-84. In such a situation, a human giving directions would've told me to take I-84 at the first exit rather than I-291 because I-84 corresponds to my intended route. It was just this kind of confusion that got me on I-384 in the first place!

Digital Object Identifier 10.1109/MCE.2016.2648019
Date of publication: 15 March 2017

106 IEEE CONSUMER ELECTRONICS MAGAZINE / APRIL 2017

The degree of similarity can vary greatly, which is why the most effective communication is a conversation as we seek a shared understanding.

I've run into similar problems with Waze. In Newton Corner, Massachusetts, there's a segment of St. James Street that goes to the Massachusetts Turnpike. The software always tells me to go to the turnpike and only after that, with a few feet to spare, does it tell me to get in the left lane to continue onto St. James Street. These problems can be solved one by one, but that's not the way our understanding works. The approach itself is not new. The Cyc project took the approach of accumulating knowledge.

The basic problem is that we are treating these programmed systems as if they think the same way as humans. I call this the *new animism* in ascribing humanlike intent to inanimate objects and gods. If a tree falls in my path, I might think the tree chose to do it rather than accepting it, or I did something to deserve it, rather than as happenstance.

It's an approach that serves us well in working with other people. We can communicate because not only do we have a shared context externally, we share similar cognitive mechanisms. The degree of similarity can vary greatly, which is why the most effective communication is a

conversation as we seek a shared understanding. If the response isn't appropriate, we adapt. Something as simple as a raised eyebrow might indicate a failure to communicate. We recognize that idioms and stories may not be shared across cultures. We also diagnose other failures of shared context, such as with people on the autism spectrum.

When using a navigation program or an autosteering car, we apply the same heuristic of modeling its understanding on ours. This works partially because the software is programmed to mimic the external characteristics of human behavior. One example is that computers are often better than people at acting human. I experienced this when using Microsoft's handwriting on the tablet PC. It was able to read my own handwriting better than I could. My writing is more akin to scribbling, yet the program does a remarkable job in recognizing my intent. It must be very smart.

Or it may just put on a good act. It needn't understand very well because it uses tricks. To a degree, its lack of understanding works in its favor because it limits the space of possibilities, whereas a human can imagine many more possible meanings and so isn't as quick to jump to a conclusion. The other mechanism is to maintain a number of possibilities and a degree of ambiguity and then eliminate the choices that don't make sense in the given context.

That's not very different from how people understand. When I was in graduate school, I thought about how language works and took an operational approach as opposed to the linguists'

The car isn't driving itself. It can stay in a lane but relies on me to follow the actual route. It's a symbiotic relationship. At one point, after the car (OK, me, the driver) was led astray by the navigation directions, I found myself on I-384 and needed to get back on I-84 to get to Boston but the navigation system told me to get on I-291 because it saw “I-291” on the sign. A few feet on the same path to I-291/I-84 it told me to get on I-84. In such a situation, a human giving directions would've told me to take I-84 at the first exit rather than I-291 because I-84 corresponds to my intended route. It was just this kind of confusion that got me on I-384 in the first place!

I've run into similar problems with Waze. In Newton Corner there's a segment of St. James Street which goes to the Mass Pike. The software always tells me to go to the Pike and only after that, with a few feet to spare, does it tell me to get in the left lane to continue onto St James St.

These problems can be solved one-by-one but that's not the way our understanding works. The approach itself is not new. The [Cyc](#) project took the approach of accumulating knowledge.

The basic problem is that we are treating these programmed systems as if they think the same way as humans do. I call this the *new animism* in ascribing human-like intent to inanimate objects and gods. If a tree falls in my path I might think the tree chose to do it rather than accepting it or I did something to deserve it rather than as than happenstance.

It's an approach that serves us well in working with other people. We can communicate because not only do we have a shared context externally, we share similar cognitive mechanisms. The degree of similarity can vary greatly which is why the most effective communication is a conversation as we seek a shared understanding. If the response isn't appropriate we adapt. Something as simple as a raised eyebrow might indicate a failure to communicate. We recognize that idioms and stories may not be shared across cultures. We also diagnose other failures of shared context such as with people on the autism spectrum.

When using a navigation program or an auto-steering car we apply the same heuristic of modeling its understanding on ours. This works partially because the software is programmed to mimic the external characteristics of human behavior. One tell is that they are often better than people at acting human. I experienced this when using Microsoft's handwriting on the tablet PC. It was able to read

my own handwriting better than I could. My writing is more akin to scribbling yet the program does a remarkable job in recognizing my intent. It must be very smart.

Or it just puts on a good act. It needn't understand very well because it uses tricks. To a degree its lack of understanding works in its favor because it limits the space of possibilities whereas a human can imagine many more possible meanings and thus isn't as quick to jump to a conclusion. The other mechanism is to maintain a number of possibilities and maintain a degree of ambiguity and then eliminate the choices that don't make sense in the given context.

That's not very different from how people understand. When I was in graduate school I thought about how language works and took an operational approach as opposed to the linguists' formal grammars. The approach of parsing a sentence into deep structures didn't make sense to me because it meant eliminating the essential ambiguity. It seemed to make more sense to maintain that ambiguity as long as possible and then eliminate possibilities that didn't work.

(This works best when words (or phrases) serve as the brain's internal representation and conversations between people are similar to our internal dialog. But that's another topic).

Computers "think" differently from people. Note that I put "think" in quotes to distinguish it from the way people typically conflate "thinking" with self-awareness. I'm just using think in the sense of cognitive process. This is part of the confusion caused by using anthropomorphic terminology for dynamic systems. We use such terms because, as we've seen, the emergent properties are indeed similar to what we see in other people. But it makes it all-too-easy to slip into projecting human cares on inanimate objects.

If we accept that these are alien intelligences, we can start to speak to them in something akin to a native language. We can designate roads as auto-steering friendly. On those roads, human drivers may see their line gradually disappear and understand they are supposed to merge while yielding. That same highway can have "Tesla" markers showing the merge path and giving the merge rules.

Driving the Tesla, I was acutely aware of the limits of the software and tested my understanding. My wife didn't always appreciate those tests or the Tesla's "judgment". It tends to drive more closely to side obstacles than people are comfortable with. The software should be as concerned about people's comfort zone as it is about safety. It had difficulty with many situations that are obvious to people, like not staying in a lane that is disappearing.

Given how intelligent the car seemed to be, why didn't it deal with traffic lanes and follow directions to turn? That's probably coming but city streets will remain a challenge especially for roads that don't have definite lanes and a myriad of other such ambiguities and unstated assumptions.

Having two different intelligences sharing the same highway can be a challenge. The onus, to a degree, is on the aliens partially because they are visitors but also because they are, for now, expecting self-driving cars to be more responsible and to yield to their human counterparts' careless behavior. If I know a car is on auto-steering I might merge into its lane knowing (sometimes naively) that it will yield.

It's frustrating for the self-driving cars (or at least their programmers and users) in that it can't take advantage of its adroitness. Such cars can drive faster and closer to their design limits while making tradeoffs for fuel use. They can also cooperate with other cars on the road and coordinate along stretches of roads. A two-lane road for such cars can have the capacity of a three-lane road; even one lane roads can be "bonded" the way network wires can.

It's understandable that the current emphasis is on solving the problem entirely within the car. Once we are able to get past emulating people we can take advantage of the real capabilities of these alien intelligences. We can also start to rethink transportations and what cars are.

When speaking to someone who doesn't speak your language you don't shout, you try to find a common vocabulary. Instead of treating cars as your best friend we need to think in terms of the new possibilities of our "smart" devices. Instead of a driverless car we might provide rides using whatever vehicle is available and appropriate for the task. Self-driving taxis rather than cars.

We need to be wary about extrapolating our love affair with these automatons as our devices become increasingly capable and intelligent. The algorithms they are built on don't really care but only act as imperfect and, often buggy reflections of what we teach them and show them. But we can't always extrapolate what happens as these systems evolve. Algorithms that work in the small may not work in the large and algorithms which apply to populations may be perverse when applied to individuals. They simply cannot care.

One positive result of trying out the Tesla's auto-steering system is that my wife now has a high opinion of my driving, or, at least, that I am better than your average Tesla.

Looking ahead

The drive assist capabilities are seen as a prelude to fully self-driving cars as well as the larger promise of “big data”. The assumption is that the more data we have the better we can see the future. Perhaps 2016 will remind people of the limits of such a view. Even if we can see the future others can too and can game it.

We also need to remember that driving is a social activity. If a self-driving car has to obey what happens when there is a posted limit of 25 MPH and humans know it's more of a suggestion but find themselves behind a car that is obliged to follow the letter of the law – will rage against the machine be the new road rage?