

A.I - the context problem.

Level: Intermediate





Context is a major problem when designing artificial intelligence systems. We recently came across a great hypothetical example.

You and your A.I powered helper robot are late to catch a plane home. You arrive at the airport and remember that you need a gift for your loved one, you also fancy a coffee.

You send your robot to the coffee shop whilst you go in search of a suitable gift. Your robot has instructions to bring you a coffee within five minutes. In the coffee shop the robot finds a long queue. Does it kill everyone in queue so it can get to the front and acquire your coffee within the time parameter you set? No of course not! It knows not to kill people so does it just push everyone out of the way? No of course not! It knows to treat humans with respect.

There are several helper robots in the queue so does it just destroy those making the queue shorter?



This example seems very stupid, of course your helper robot knows not to kill or harm people. It also knows not to damage other people's property. You just want a coffee, not death and destruction.

But how does your helper robot know these things? As a human it all seems instinctive, you know right from wrong, you understand the context of the situation and the circumstances under which the command to buy a coffee was given. You learnt a lot of this stuff as a child.

However, to avoid an awkward situation in the coffee shop which results in you being arrested, the information that gives your helper robot context is either installed when its mind was switched on, or it learns it (hopefully not by killing coffee drinkers to discover if that is right or wrong).

It is suggested that the average person makes about 35,000 choices per day. All of those choices are subject to context. We accept that most of those choices will be subject to a common set of rules, but the context could be slightly different or completely different each time.



An A.I acting within the wrong context is not so bad if it is buying and selling shares or creating a work of art. However, as A.I becomes more integrated into society it will be given the ability to move around, enabling it to interact with us physically.

Self-driving cars do this already and to date, their safety record has been pretty good though people have still died as a result of using the technology.

The biggest problems will surface as we take A.I into battle with us. When the A.I has the ability to kill the enemy who gives the kill order? One way to try and solve this problem is to simulate battles and to train the A.I in what we might see as a computer game. Another way is to place a human in-between the robot and the decision to kill.

Currently “Robot Wars” is an entertaining TV show but one day robot wars will be fought alongside our young men and women, people will die at the hands of a machine.



There will not be Terminators roaming the streets, we will not wage a war against the machines. However, we will wage war using intelligent machines. In fact we do that ready if you consider the ability of smart missiles hitting a target a meter square that is a hundred miles away.

We believe that A.I is a wonderful thing, the greatest and most powerful tool that mankind has ever created. It is a force for good.

There will be a few mistakes on our A.I journey but we will solve the context problem.

Coffee anyone?



Created for you by Intelligent Systems Development Alliance Ltd.
A non-profit funded by donations, sponsorship and commercial operations.
Please join us, go to: www.isda.world