Journal of Geoethical Nanotechnology
Vol. 10, Iss. 1 (2020) 15 - 25

Terasem Movement, Inc.
www.TerasemCentral.org

NANOROBOT ETHICS

IDO BACHELET

Ph.D., Faculty of Medicine and Life Sciences, Bar Ilan University, Israel.
ido.bachelet@biu.ac.il

Together with students across varying disciplines at the Bar Ilan University, Dr. Ido Bachelet works with molecular sized robots to automate processes within our biological systems to repair and improve one's overall health.

Keywords: Nanobot, nanotechnology, nanoscale, Kurzweil, Merkle, Drexler, Rothemund, Seeman, DNA origami, Gemtuzumab, Asimov, ethics, Kinect, Emotiv, Xboa, EEG, Bar Ilan University.

Nanorobot[1] Ethics is a certain branch of nanotechnology that I've been working on for the past several years. What we are currently working on is how to encode ethics, or the way we understand ethics, in molecules so nanoscale machines we're making can use them.

Nanorobots have been envisioned for the last thirty years or so by people like Ray Kurzweil[2], Ralph Merkle[3] and Eric Drexler[4]; nanorobots are amazing because they can do many things. For example, they can automate our own biology; they can automate processes. Imbalances in molecules that cause disease can be repaired by nanorobots. They can for example, perform surgery on their own without the need to invasively cut someone open. Another example is the virtual control of drugs. Even after four decades of remarkable progress in drug delivery systems, we still cannot control drugs. By control, I mean the way we use computers to control processes and machines in the real world. For example, when we say that a drug is too toxic to use, this is not a drawback of the drug, it's basically a drawback of ourselves because it means we don't know how to control that drug properly.

---

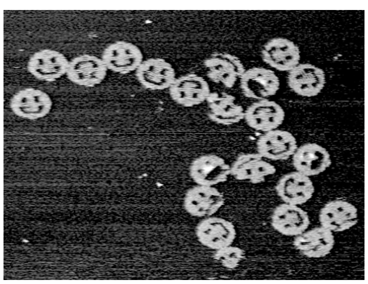
[1] Nanorobotics - molecules with a unique property that enables them to be programmed to carry out a specific task. Berger, M. (n.d.) What are nanorobots? *nanowerk*. Retrieved from https://www.nanowerk.com/what-are-nanobots.php
[2] Ray Kurzweil – one of the world's leading inventors, thinkers, and futurists, with a 30-year track record of accurate predictions. Retrieved from http://www.kurzweilai.net/ray-kurzweil-biography
[3] Ralph Merkle – Senior Research Fellow at the Institute for Molecular Manufacturing, Co-founder of Nanofactory Collaboration and Chair Emeritus, Nanotechnology at Singularity University. Retrieved from http://www.merkle.com/
[4] Eric Drexler - Often described as "the founding father of nanotechnology". Retrieved from http://e-drexler.com/p/04/04/0404drexlerBio.html

If we only had a computer or a robot that we could link this drug to, we could solve that problem, but obviously for the last thirty years we've been unable to imagine how these nanorobots should look. I think the reason is basically in our minds… whenever we hear the word "robots", that's what we think about, right? Which reminds me again how icons limit our minds, limit the way we think.

Remember what this is? This is what phones used to look like this. My four-year-old son doesn't know what that is anymore. But still, when you go to Toys R Us, they sell rotary phones, even though they are not in use anymore. This is how strong the power of icons is in our minds.

Credit: Public domain

To build our nanorobots we're using another technique, which we call DNA origami[5]. DNA Origami is a method invented by Paul Paul Rothemund[6] from Cal Tech about ten years ago and originated in the works of Nad Seeman[7] of NYU even thirty years ago. This is a new technology enabling us to take a single piece of DNA, the same DNA that is in our cells or in the cells of any other organism in nature, except for RNA and viruses, of course, that don't use DNA, they use RNA. We can take any piece of DNA and program it's folding into virtually any 3D or 2D shape you want. For example, these smiley faces are about 100 nanometers in size.



For example, these smiley faces are about 100 nanometers in size. Only 100 nanometers and we can make trillions of them in a tube in a very simple reaction. So we're using DNA Origami to build nanorobots. What you see below is nanorobots; this is what they look like. The cubes you see here are actually machines, tiny machines, the size of each one is 15 nanometers and we can program them; they have moving parts, they have screens and axes all made of DNA molecules. We published that in *Science*[8] and I think it's an open source article so you can catch that on Google.

These robots are great, they perform one basic task; you can load these robots with any combination of cancer drug(s) that you want and the robots can actually control the drug for you.

---

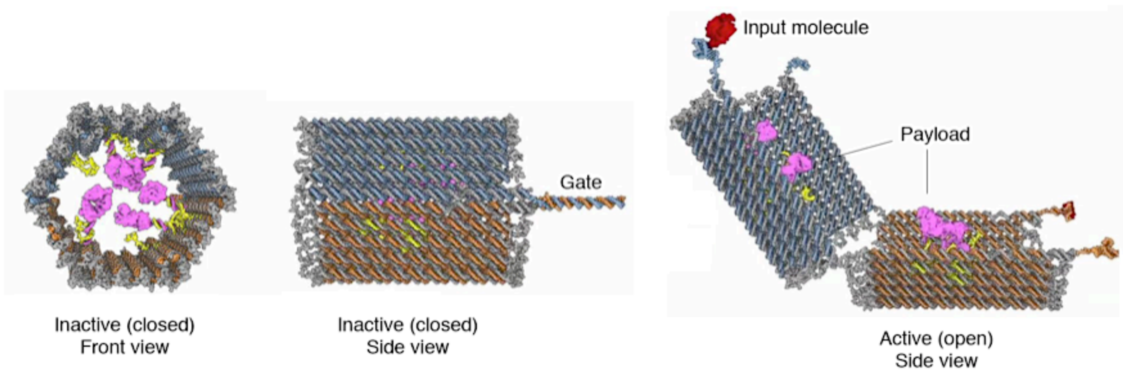
[5] DNA origami - the process in which long single-stranded DNA molecules are folded into arbitrary planar nanostructures with the aid of many short staple strands. Kuzuya, A. & Komiyama, M. (2010). DNA origami: Fold, Stick, and Beyond. Royal Society of Chemistry: *Nanoscale.* Retrieved from ttps://pubs.rsc.org/en/content/articlelanding/2010/nr/b9nr00246d#!divAbstract

[6] Paul W.K. Rothemund, MS – Research Professor at Caltech working on integration of DNA origami into microfabricated devices. Retrieved from http://www.dna.caltech.edu/~pwkr/

[7] Nadrian C. Seeman, PhD is a Professor of Chemistry at New York University in New York City, NY. Retrieved from https://as.nyu.edu/content/nyu-as/as/faculty/nadrian-seeman.html

[8] Douglas, S., Bachelet, I. and Church, G. (17 Feb 12). A Logic-Gated Nanorobot for targeted Transport of Molecular Payloads. *Science*, Vol. 335, Issue 6070, pp. 831-834. Retrieved from http://science.sciencemag.org/content/335/6070/831

The robot looks like a tiny clamshell in the shape of a hexagon. What you see inside is these yellow stems, which are in fact loading sites on which you can load anything you want starting with small molecules, drugs, or proteins. The pink blobs you see there are special therapeutic antibodies for cancer. This was a drug called Gemtuzumab[9], which was taken off the market for a while for excessive toxicity. You can take the same drug, put it in the robot and the robot can turn it on and off basically just by opening and closing as you can see here. What's causing the robot to open or close is a combination of molecules in the environment, which we program the robots to look for.



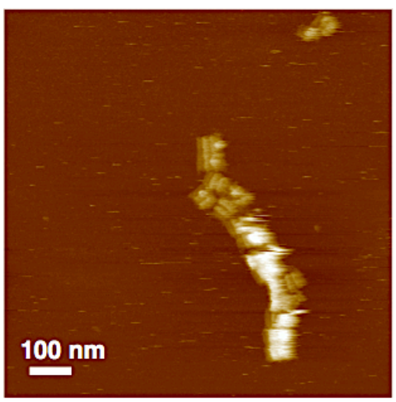
For example, it can identify the molecules on a cancer cell, it can identify molecules released from certain sites, they can identify certain tissues, certain organs, and we have a large target bank we're expanding every week. We already have robots that can identify about twelve types of cancer cells and can activate cancer drugs, drugs that have been considered too toxic to be used, and they can activate these drugs, turning them on and off. They don't release the drug; the drug is connected to the robot, but the robot closes and opens, turning the drug on or off as needed when it identifies the target cell or condition. We have robots that can identify nerve cells that fire excessively, for example in pain and epilepsy, and we are working on robots for diabetes which will be able to respond to glucose levels in the serum and plasma and activate or deactivate insulin.

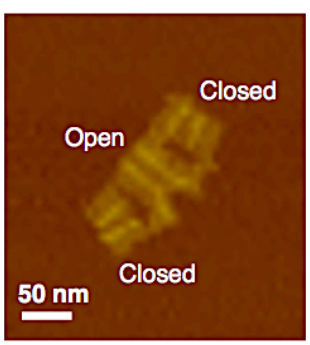
One final note about that, we've built them such that they are not immunogenic, they don't release any immune response and we can tune their survival in the blood for up to three days. These are the bots, but now before I move on to the ethics part, one of our major problems is that the targets of our drugs, which can be two or more cells, bacteria, viruses or fungi, these targets keep undergoing evolution.

[9] Nelson, R. (12 Jun 10). Gemtuzumab Voluntarily Withdrawn From US Market. *Medscape Medical News*. Retrieved from https://www.medscape.com/viewarticle/723957

They keep getting better and better every single minute and we don't, so a generation for us is twenty years, a generation for them can be twenty minutes. Every twenty minutes they have a chance to outsmart us - and single robots cannot outperform this, but one of the advantages of swarms, groups or collectives in nature is that they can adapt to changes - so if we could program the robots, program these DNA origami nanobots to behave as a swarm, we could have a chance of adapting to drug resistance in tumor cells or bacteria and we could possibly outsmart them. To do this we are learning from natural swarms like this swarm of birds this photo was taken over England. The basic idea about swarms is that they exhibit very elaborate behaviors that emerge from very simple interactions between individual agents, which can be individual birds, fish or bacteria. We're learning how these behave and we are copycatting this into our nanorobot system. We're establishing communication between robots and we see the behaviors that emerge. These are some examples of behaviors we are already capable of programming.

Just like you see these insects below, forming physical bridges between parts of the same tree or leaves, we are able to program nanorobots to do the same so they can connect two points like A and B, which for example, can be like cells. This is good to support and guide the re-growth of cells from point A to point B across an injured tissue, like an injured spinal cord. These robots can form a bridge between two halves of spinal cord so that each neuron, each nerve cell, can find its own half across the gap, across the injury.

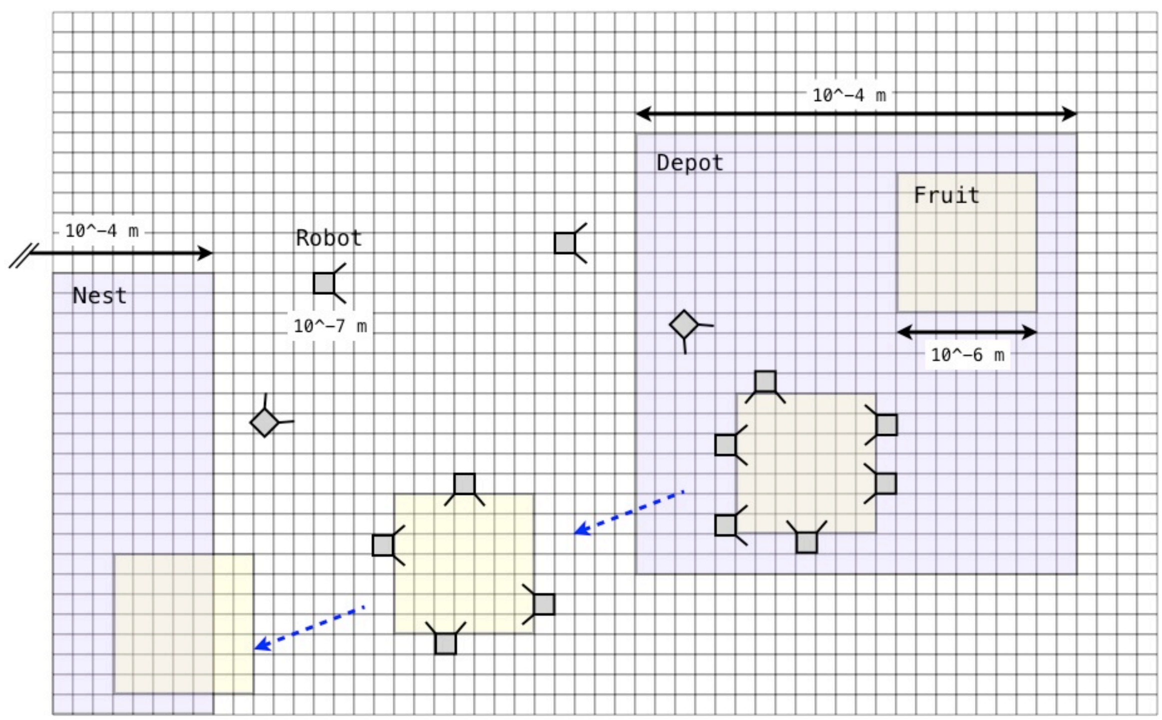




In the lower part you see robots that communicate physically between each other and these can regulate their activity such that when one of them is open, as you can see in the middle, the other ones are closed. This is great because we were able to use this as a tool to administer drug combinations. In drug combinations often you take many drugs in parallel and they induce side effects. But here you can program a cue, a sequence that if robot A is active and the other ones are inactive, then A becomes inactive and B turns on and so on and so forth. So several drugs can operate in parallel without actually combining with each other.

Another example of something we're working on is robots that can fetch something to a nest. They can emanate from a single point which we call "nest", they forage for food, again the food can be molecules or viruses for example, and the robots can actually forage and fetch that to a certain point which is more accessible to us and we can, for example, surgically remove things or similar.

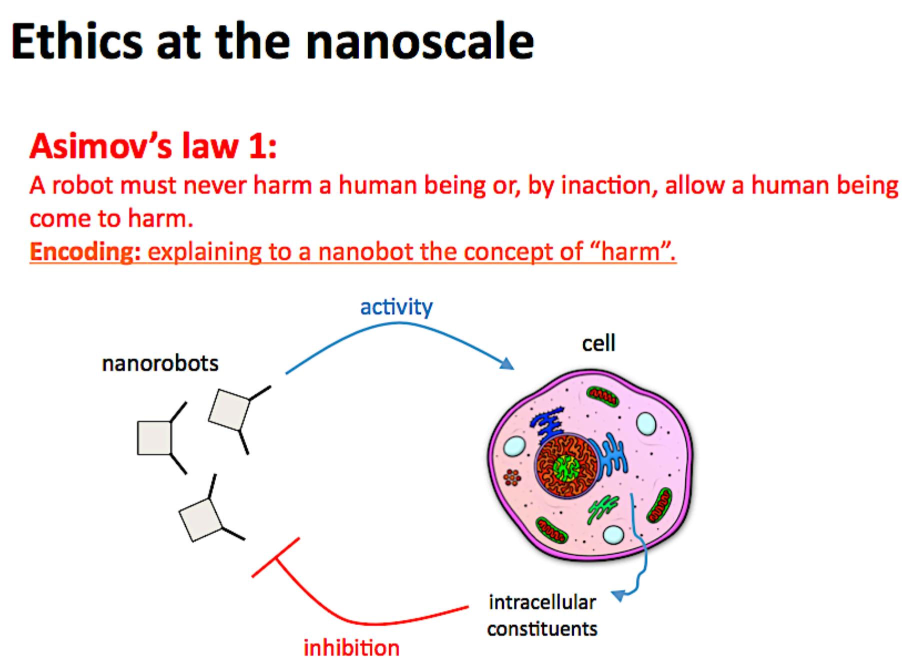
The last example to show is robots that can interact in the very basic way, like transistors interact in your computer. The processor of the computer is just transistors interacting in certain ways to generate logic gates, and logic gates interact in certain ways to do computation. So we program robots, which you can see on the figure on the top right,



we programmed them to interact in special ways that mimic the ways that transistors interact in your home computer. Using this approach, we were able to program nanobots to emulate logic gates. As you can see in the graph at the bottom these are robots carrying drugs and the output, meaning the drug activity, is actually emulating logic gates. These logic gates again are the basic components of your computer and we can cascade those robots so we can connect two of these robots so that their output is the input of the third system. What we have is we already have systems built from cascaded robots that can perform computations of 8-bit and 16-bit (by the way we do this in living animals; we use living insects and the robots write their outputs of drugs activated on the insect itself).

Now we get to the interesting part. These robots can, say anytime within the next few years next year or so, be integrated into medical procedures, into therapeutic procedures and they can integrate into our society just in the basic way that Isaac Asimov[10] envisioned in the early and mid Twentieth Century, when he and others wrote their novels about intelligent robots. The question is: How do we make sure these robots do not harm human beings? How do we make sure they are safe? Isaac Asimov framed these very famous three laws of robots, which supposedly every robot, for integration into our society should be programmed to follow.

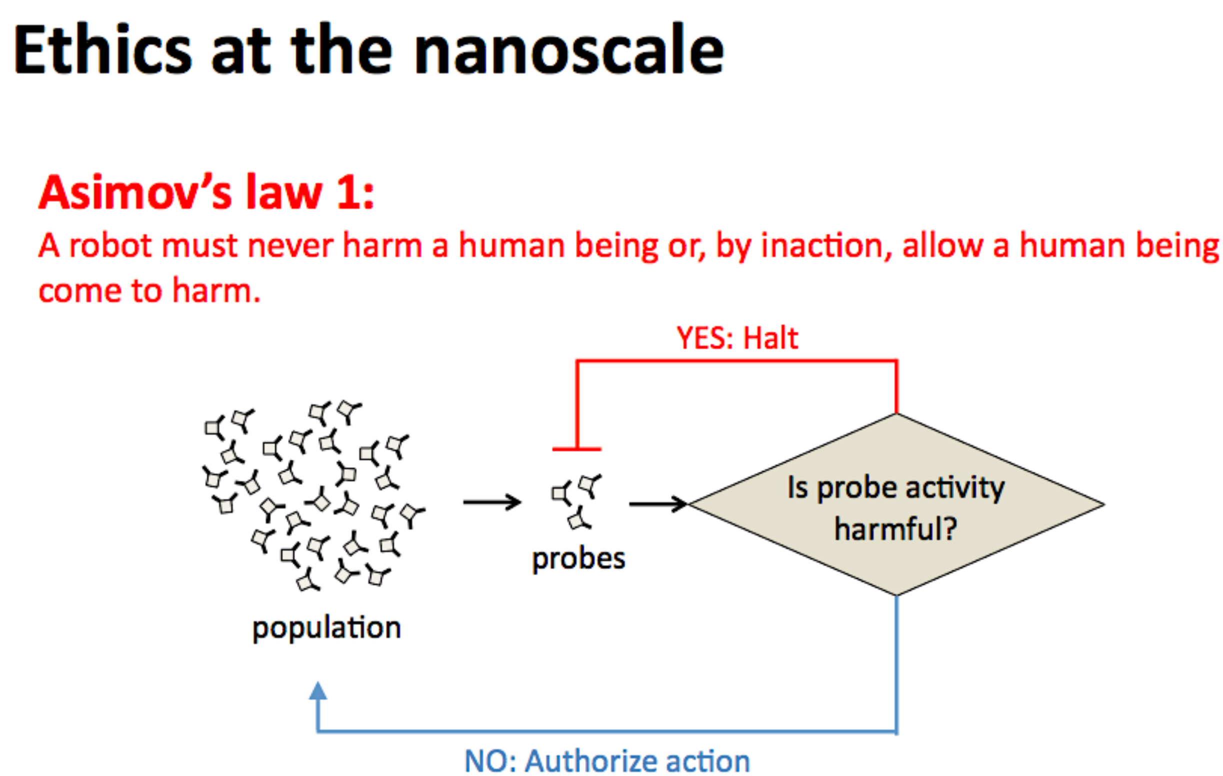
The first rule, Law #1 means that the robot must never harm a human being or by inaction allow a human being come to harm. Law #2 is a robot must always obey human commands, except when these are in contradiction to Law #1. And Law #3 is a robot must always protect itself, unless when this contradicts laws 1 and 2. Later, another law was added, Law #0, meaning hierarchically it is higher than the other three, that a robot must never harm humanity.



[10] Isaac Asimov - the quintessential author, who in his lifetime wrote over 500 books that enlightened, entertained, and spanned the realm of human knowledge. Retrieved from http://www.asimovonline.com/asimov_home_page.html

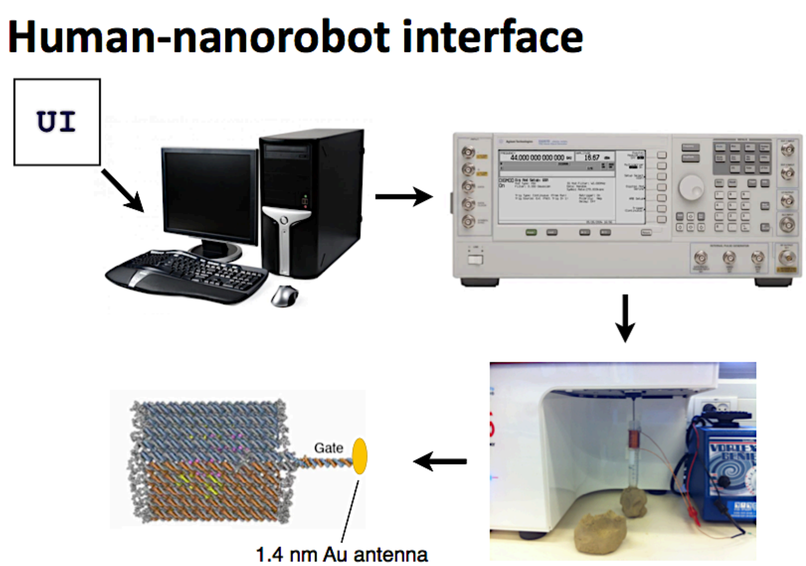
This is more global, but we manage to encode Law #1 in molecules and also generated a very unexpected and very nice application of nanobots as well. For the robot to be able not to do harm, first it must understand what harm is. So, the first question of the first law is: how do we explain to a robot what harm is, what damage is? If you think about a big robot, like a macro-scale robot, the Roomba that cleans your floor for example, it's very hard to explain to those robots "what harm is". What does it mean when a human being gets hurt?

On the molecular scale this is very easy. All you need to do is find a molecule that is normally inside cells and tell the robot to look for this molecule outside of cells. Whenever the robot finds this molecule outside of cells it interprets this as damage. Meaning the cell exploded and was damaged somehow, allowing the molecule to spill outside, so normally it should only be inside, but now it's also outside. So the robots understand they do damage right now, a very easy concept to explain in molecules. What we technically do is design DNA molecules that sort of assume a structural confirmation only in the presence of that molecule, but not otherwise; we integrate those molecules as gates in the robots. Using this approach, populations of robots can send probe robots that activate their payload or activate their drug on the target cell. Now they ask the question, is the probe activity harmful? If yes, if the molecules spill from the inside to the outside of the cell and the DNA sequence that we designed identified it and assumed a second confirmation, then it means the robot should be locked, that is what the DNA is designed to do. If not, then the robots continue as usual and the entire group can do this.



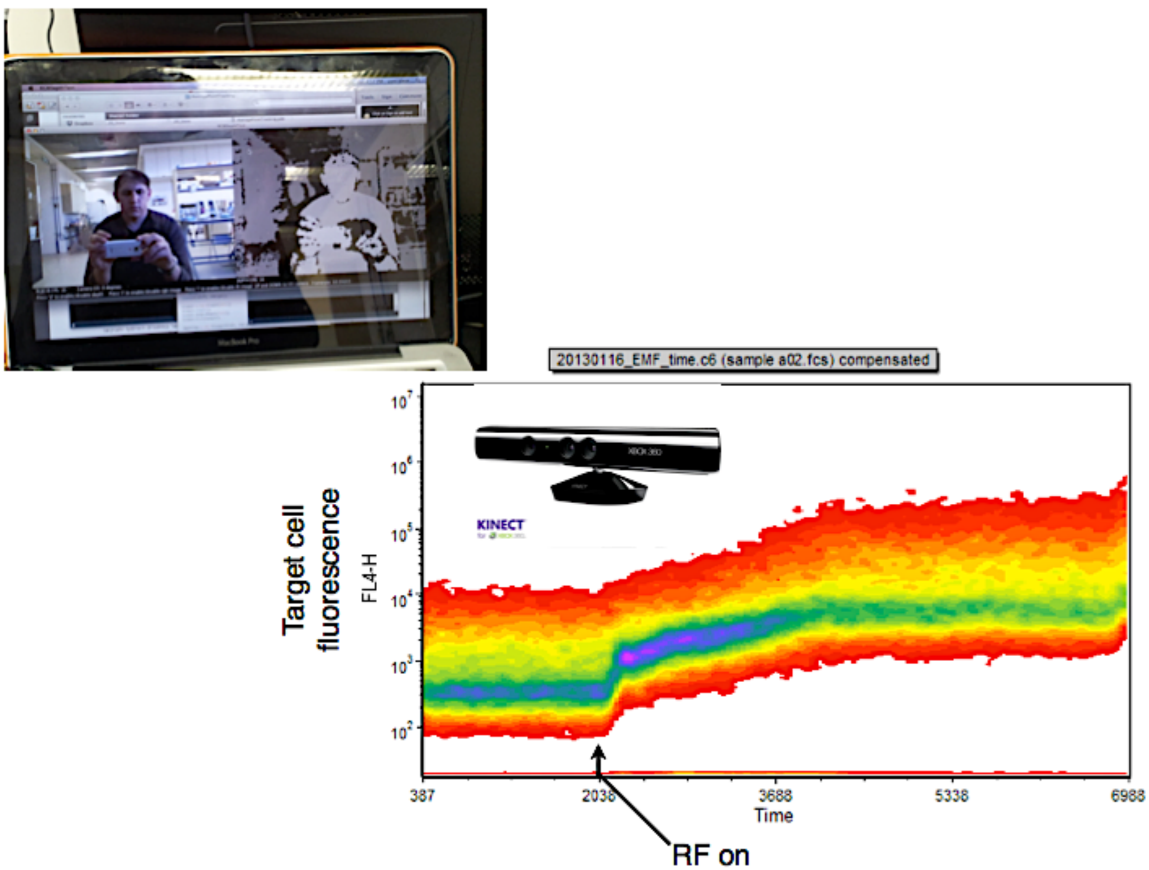
One thing that we discovered we can do is explain to the robot what harm is. This enables us to monitor the damage to make sure we are effective. We can make sure we are safe, but we can make sure that when we do damage, we are effective. When do we want to do damage? When we are treating cancer - so one of the major problems in cancer, as you all know, is that cancer cells often develop resistances to the drugs we give. We took a group of robots like the one you see here and we loaded the group with four different types of cancer drugs. Now initially the group only activates drug A. All the while the robots monitor the damage that drug A does and they are actually looking for that molecule to spill outside of the cancer cells and it means that they are succeeding at killing cancer, but once they stop sensing that molecule, once the cancer cell stops secreting that molecule, it means the drug A is no longer effective. At that very moment the robots shut off drug A and turn on drug B. They can sort of chase the tumor throughout the mutations it tries to employ in order to develop resistance to those drugs. We can sort of attack the tumor again and again and adapt to the changes and outsmart it. We already use this successfully to show that a very aggressive tumor, in vitro, can be eliminated without giving it a chance to develop resistance.

So these are the questions we face: Can human intentions be encoded in molecules? Can harm to humanity be encoded in molecules? Can these be solved at the level of human-nanorobot interface? These apply to Isaac Asimov's Laws #2, #3 and #0 because Law #1 was pretty simple to encode in DNA, but Laws #0, #2 and #3 are harder. For this we're also working on the interface between the DNA origami robots and human beings and I'll show you in these two final graphics what we did with that. Instead of teaching the robots to look for molecules, we built versions of the DNA robots that carry antenna. Now the antenna (as you can see here), are made of gold and they are very small, only 1.5 nanometers in length, but they're still antennae in the sense that when we apply an external electromagnetic field the antenna heats up and that heat is sufficient to melt the gate and open the robot transiently. When we turn off the magnetic field the robot turns off, the robot closes up again; an Xbox can actually control these robots and this is what we did.



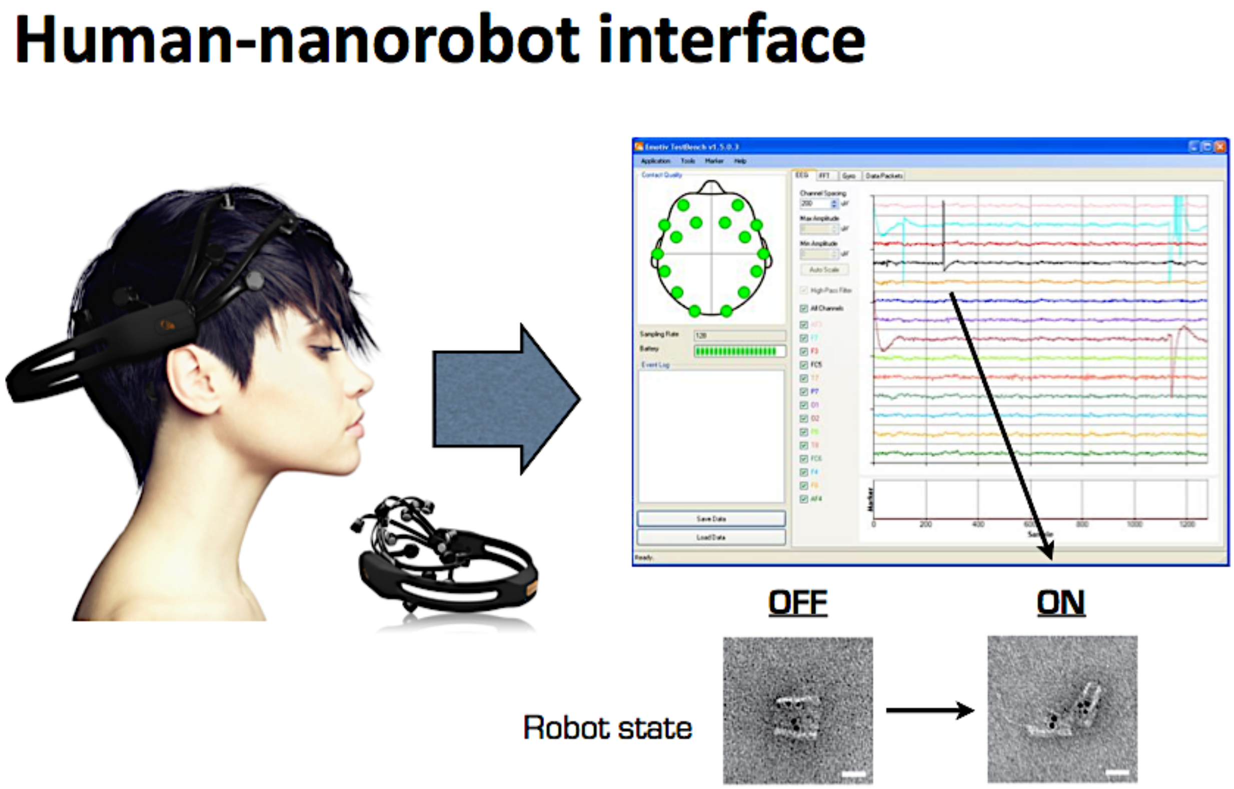
Here you can see one of my students writing an algorithm, a software program that enables us to control these DNA nanorobots, these molecular robots, just by a joystick or by a gesture using a motion sensor like the Kinect[11]. At the bottom you can see cancer cells, the graph represents cancer cells that at the middle of the graph just by a hand gesture are received by the Kinect. We actually open the robots and the robots carrying a combination of cancer drugs actually started killing the cancer cells - you can see that after forty-eight hours they all die. Again, this is controlled using just the Xbox.

Human-nanorobot interface



[11] Microsoft Kinect – The Kinect sensor incorporates several advanced sensing hardware. Most notably, it contains a depth sensor, a color camera, and a four-microphone array that provide full-body 3D motion capture, facial recognition, and voice recognition capabilities.
Zhang, Z. (27 Apr 12). Microsoft Kinect. *IEEE MultiMedia*, Volume 19, Issue 2.

The final example I show is robots that can be controlled by thought. We do that using a neuro headset, the Emotiv[12] you can find on-line. What we do is we are trying to look for EEG[13] patterns that are generated in response to cognitive load. Cognitive load is any pattern, not only cognitive load but also bipolar disorders, schizophrenia and other diseases where we can't always find a certain molecule that the robots can look for. The idea is to write algorithms and we have students doing that right now, we're writing algorithms that connect between EEG patterns we pick up with the Emotive to electromagnetic fields that can open or close the robots; we're still working on that.



[12] Emotiv - a portable, high resolution, 14-channel, EEG system. It was designed to be quick and easy to fit and take measurements in practical research applications. Retrieved from https://www.emotiv.com/epoc/?gclid=Cj0KCQiA6JjgBRDbARIsANfu58F44UQZr_ovD7eQhl6_t_g1UTYpoGflRGDp6pXMHXPkCC4F3dEnRf4aAuMnEALw_wcB

[13] EEG - a test that detects abnormalities in your brain waves, or in the electrical activity of your brain. Electroencephalogram (EEG): What is an EEG? Johns Hopkins University School of Medicine Health Library. Retrieved from https://www.hopkinsmedicine.org/healthlibrary/test_procedures/neurological/electroencephalogram_eeg_92,p07655

The final lesson is that encoding human ethics in nanorobotic systems requires a multidisciplinary approach that can be partially solved just at the level of molecules to be designed, but some of it has to remain at the human molecular interface and this is something that again links back to us, meaning we can control the ethics of the robots we make. This is something that we are still thinking about and it is still totally open to discussion.

I want to thank my amazing team of students at Bar Ilan University[14] who work on these nanobots and try to come up with solutions to those nanorobot ethics problems. I'm working to build a group that is multidisciplinary so it has students from biology, physics, computer science, architectural design and electrical engineering as I feel this is the most proper way of addressing these challenges.

References

Assimov, I. (n.d.) Assimovonline. http://www.asimovonline.com/asimov_home_page.html
Bar Ilan University (2018). Ramat-Gan, Israel. Retrieved from https://www1.biu.ac.il/en-about
Berger, M. (n.d.) What are nanorobots? *nanowerk*. https://www.nanowerk.com/what-are-nanobots.php
Douglas, S., Bachelet, I. and Church, G. (17 Feb 12). A Logic-Gated Nanorobot for targeted Transport of Molecular Payloads. *Science*, Vol. 335, Issue 6070, pp. 831-834. Retrieved from http://science.sciencemag.org/content/335/6070/831
Drexler, E. (2009). http://e-drexler.com/p/04/04/0404drexlerBio.html
Emotiv (2018). Retrieved from https://www.emotiv.com/epoc/?gclid=Cj0KCQiA6JjgBRDbARIsANfu58F44UQZr_ovD7eQhl6_t_g1UTYpoGflRGDp6pXMHXPkCC4F3dEnRf4aAuMnEALw_wcB
Johns Hopkins University School of Medicine Health Library. Retrieved from https://www.hopkinsmedicine.org/healthlibrary/test_procedures/neurological/electroencephalogram_eeg_92,p07655
Kurzweil Network. (2018). *Accelerating Intelligence*. http://www.kurzweilai.net/ray-kurzweil-biography
Kuzuya, A. & Komiyama, M. (2010). DNA origami: Fold, Stick, and Beyond. Royal Society of Chemistry: *Nanoscale.*
Merkle, R. (n.d.). http://www.merkle.com/
Nelson, R. (12 Jun 10). Gemtuzumab Voluntarily Withdrawn From US Market. *Medscape Medical News*.
Rothemund, P.W.K. (12 Nov 18). http://www.dna.caltech.edu/~pwkr/
Seeman, N.C. (n.d.) Professor of Chemistry at New York University, New York City, NY. Retrieved from https://as.nyu.edu/content/nyu-as/as/faculty/nadrian-seeman.html
Zhang, Z. (27 Apr 12). Microsoft Kinect. *IEEE MultiMedia*, Volume 19, Issue 2.

[14] Bar Ilan University (2018). Ramat-Gan, Israel. Retrieved from https://www1.biu.ac.il/en-about