Unleashing Your Inner Maker

DR. DANIELA RUS

Andrew (1956) and Erna Viterbi Professor of Electrical Engineering and Computer Science and Director of the Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT

The President's Distinguished Lecture Series • Stevens Institute of Technology

November 28, 2018

Thank you so much for this generous and kind introduction. Good afternoon, ladies and gentlemen. I'm delighted to be here. It is a privilege and great honor. I am especially glad to be here for the launch of your new AI center. This center is visionary, and I have to say I love its name. I read it as "See AI," so clearly, Stevens Institute sees the potential of AI.

Now I want to start out with a story and let you know that when I was a kid, I used to love the books of Jules Verne and the movies of Jacques Cousteau. Some years later, after I learned how to make and program robots, I met a student who also loved the books of Jules Verne and the movies of Jacques Cousteau, and we decided to make a robot together, our own version of the undersea observatory. We called this robot AMOUR, and AMOUR stood for autonomous modular optical underwater robot, but making this robot was really about love. The robot was elegant, it moved beautifully in water, it took beautiful pictures, which could be transmitted through an HDTV-quality optical modem, all enabled by innovative techniques.

Now on a very hot, hot summer day off the cost of Singapore, we wanted to take this robot for a ride to see just how far it could go. We were feeling very confident that day, so we decided to disable all the security features of the robot to give it more juice, and then off the robot went. Now I should tell you that this robot was our only copy and it had taken us three years to build it. So we waited and waited and waited, but the robot never came back. We were devastated.

Now the thing about science is that great failures teach you great lessons. It is also true that the first copy of an invention takes a long time to make, but subsequent copies are much faster. So equipped with this wisdom, we went back to Boston and we grounded ourselves. Three months later we had three improved versions of AMOUR. We took the new robot to sea, and we watched it move, and as it performed flips, we started wondering, what might the heirs of AMOUR look like? Maybe softer, gentler, more agile.

Please meet our soft robot fish called SoFi. This robot fish is very special. It moves back and forth by undulating its tail. The fish is very agile. It can change the swimming direction in about 100 milliseconds, which is the same as real fish. This maneuver allows little fish to escape the big fish who want to eat them. We took SoFi to the ocean to see it swim. Here you can see me next to the fish. You can also see SoFi's big black eye — a fisheye lens camera. And just look at SoFi swimming elegantly past beautiful corals and into a beautiful cabbage patch; it saw many schools of fish with its fisheye lens camera, and it even found Nemo.

> **"...the thing about science is that great failures teach you great lessons. It is also true that the first copy of an invention takes a long time to make, but subsequent copies are much faster."**

Okay, so it took us three years to make AMOUR, three months to make three copies of AMOUR, and today in the lab we can make SoFi in a few days. This rapid change of pace in technology may sound familiar to many of you, because just two decades ago, computing was a task reserved for the expert few; computers were large and expensive and you needed real expertise in order to know what to do with them. But today computing is so normal, we don't even notice how we depend on it. This raises an interesting question: In a world so changed by computation that helps us with many compute tasks, what might it look like with robots helping us with physical tasks and AI systems helping us with cognitive tasks?

How much work will we be able to offload to machines to save lives and increase the quality of our lives? Now I believe that autonomous driving will absolutely ensure that there will be no road fatalities, it will give our parents and grandparents a higher quality of life in their retirement, and it will give us all the ability to go anywhere anytime. So please indulge me and imagine driving home in your autonomous car knowing that it has the smarts to keep you safe and the intelligence to make that ride fun. You have to bring home supplies for dinner, so you pull up at the grocery store, where your car connects with your refrigerator, which then communicates with the store and they figure out what items you're missing for tonight's menu, and a few minutes later a robot brings you a box with all the items you need.

And when you get home, you pass the box to your kitchen robot, and you might even let your children help with dinner because even though they make a mess, the cleaning robot will help clean up. Now, some of you may be thinking that this sounds like one of those cartoons about the future that never comes to pass, but I would like to tell you that that future is not that far off. Today robots already work side by side with people in industrial and in domestic settings. They help surgeons in hospitals, they help pack goods in factories, they vacuum our floors, they cut our lawns, they even milk our cows. And in the future they will do so much more for us.

But in order to get to this future, we need to solve a challenge: How do we make and code intelligent robots? And to be precise, a robot has a body and a brain. In order for the robot to be able to do any task, that robot has to have the body capable of doing that task and the brain capable of controlling the body to do the task. A wheel-based robot will not climb stairs. Today, the vast majority of robots are inspired by this one and look a little bit like this robot manipulator called the Unimate, a robot that was built in 1961.

In 1961, we had one robot — the Unimate. By 2020 we are projected to have 31 million industrial robots. These industrial robots are masterpieces of engineering that can do so much for us. Yet these robots remain isolated from people on factory floors because they are big and heavy and dangerous to be around. But a world with pervasive robots helping us with a variety of tasks needs a different kind of machine, needs a machine that is safe to be around, that is intelligent, that understands the people around it. Now, by comparison to industrial robots, nature does things very differently, much more dexterously and much more intelligently. Just look at this octopus to see how agile it is squeezing and escaping through a hole. Or look at the elephant, whose trunk is able to move with great delicacy to pick up the banana from the hand of the child and feed itself, and yet that trunk can also be used to fend off contenders.

> **"But in order to get to this future, we need to solve a challenge: How do we make and code intelligent robots?"**

So I believe that we need to develop the technologies that help our machines look and act more like natural systems, and in order to do that, in order to get to the full potential of robots, we need to advance the science of autonomy. We need to create better machine brains, better bodies and better interactions. Now, in order to achieve these three challenges, progress is happening simultaneously in three fields, and I just want to define them so we are all on the same page.

On one hand we have robotics. Robotics puts computing in motion and gives machines the ability to move. AI gives machines the ability to reason, enables machines to hear, to communicate, to interact with people, so it gives machines human-like qualities. And machine learning cuts across both AI and robotics. Machine learning today looks at data and aims to learn a pattern that explains the data and makes a future prediction. Now, today, machines have limited ability to reason, but through the use of machine learning enabled by the vast amounts of data, we are achieving extraordinary progress. All fields that have a lot of data can benefit. I want to show you an example because it really illustrates the importance and the scope of advancements in machine learning.

"I believe that we need to develop the technologies that help our machines look and act more like natural systems, and in order to do that, in order to get to the full potential of robots, we need to advance the science of autonomy."

In a recent study, doctors and machines were given scans of lymph node cells, and they were asked to label them as cancer or not cancer. The AI system, the machine, made 7.5 percent errors compared to the humans' error at 3.5 percent, but working together, humans and the machine achieved 0.5 percent error, which is 80 percent improvement in the performance. This is extraordinary. So instead of thinking about what the machine does and what the human does, it's better to think about how the machine can be an augmenter, can be a capable tool that enables the human to do more.

I also want to show you very quickly that much of what is happening in the field of AI and machine learning today is due to a technique called *deep neural networks*. In deep neural networks, we have networks that look like this, usually millions and millions of nodes, and data that has been labeled by humans is presented to the network in order to learn the inside weights. This is done so that when a new piece of data that has not been labeled gets presented to the network, that data gets labeled correctly. For example, in this case, it's a beach.

For the case of classifying images, this process is done in two steps. There is an image segmentation process and an object recognition process. In image segmentation, if you give the network images like the ones you see in this picture, the system has to group together all the pixels that go with the same object, and so you can present pictures like this and pictures like this.

And once you have segmented the objects in the image, the next task is telling what those images represent. And do you know how this is done? Well, this is how it's done: it's done manually in call-like centers. So there is a significant manual step in the performance of today's extraordinary machine learning engines. The other thing you should remember when you think about the scope of today's machine learning systems is what it means for the machine to learn. In the context of pictures, learning means that when the machine sees, let's say, a picture of a bottle of water, and says, This is a bottle of water, that means that the pixels in that picture look like the pixels in other images that people said, Those are bottles of water.

The machine has no idea that, in fact, a bottle of water is something you use to drink water from. The machine has no idea what to do with a bottle of water. What do you do with it? Do you eat it, do you throw it, do you play with it? None of that is known to machines, but people know very well what to do with a bottle of water, and we don't need millions of images to know this. This is important to know about today's deep learning techniques. The other point is that even though the machine learning systems today have extraordinary performance, their performance is not perfect.

To illustrate this, I want to show you these two examples of images. Look at them. Do they look the same to you? They should because they look the same to me. But in fact, they are slightly different. One picture was created from the other by injecting some small noise. The noise is not perceptible to the human eye, and yet this is sufficient to trick the network that correctly recognized the first picture as a dog to recognize the second picture as an ostrich.

Finally, take a look at this picture and watch this kid. This is an 18-month-old kid that's watching this scene for the very first time, and look at what this kid does. So this kid has understood the intent of the adult, has figured out an action that solves the problem the adult has and now is making contact to see, Have I understood you correctly? We are so far from having machines that have this level of intelligence and reasoning. A truly grand challenge is to build machines capable of this kind of reasoning.

> **"So instead of thinking about what the machine does and what the human does, it's better to think about how the machine can be an augmenter, can be a capable tool that enables the human to do more."**

Now how far are we from that? So let me come to robotics and to the actual capabilities of today's machines. I would like to use as a first example transportation, because transportation is getting close to becoming a utility. I'm a very big believer in the potential for autonomous transportation. Now most solutions to autonomous driving that you read about in papers and that are promoted by companies treat driving in two phases. The driving is usually done in a closed environment. In the first phase, the robot car goes and maps every single road. The mapping is not a Google Map-like thing; it's mapping in sensor space. Then, in the second phase, the vehicle can reason how to go from point A to point B. We have a lot of robot cars that operate on that same principle. In general, one could take any car or any wheel-based vehicle and turn it into an autonomous self-driving robot by doing the following process: you take your platform, you make it drive by wire, you add sensors to it, you then create perception algorithms for mapping, obstacle detection and localization. Localization is usually done by looking at the infrastructure and in particular at the signature of the local neighborhoods. You also need a planning and control module to enable the vehicle to follow a trajectory and go around new obstacles when they are detected. So this is the autonomy recipe. Anybody can take their car and turn it into an autonomous vehicle following this pattern. So what's the problem?

Well, the problem is that, in fact, the autonomy pipeline — which starts with looking at the world, figuring out where are the obstacles, figuring out where the vehicle is located, figuring out what to do, and finally executing — the problem is that in this pipeline, all these modules have to be carefully programmed and coded for each type of road situation. In other words, if you have nighttime driving, you need special modules to treat nighttime driving. The same if you're on a road with no lane marking or with rainy weather. Can we do better? Right now these robots are quite rigid. They are not very adaptive, they cannot reason their way out of the situations that they have been explicitly programmed for.

In my lab, we are working on increasing the brain capacities and the reasoning capacities of autonomous cars, and we are looking at end-to-end learning. This means going from a single-camera image into action, in other words, a steering angle for the vehicle. Here is an execution of the end-to-end approach. This car is taking a first ride in the countryside after being trained in the city. It's going about 23 miles an hour. If I think about my first drive on a country road, it was not as smooth as this one.

In red and blue are the ground truth versus the car's predicted trajectory. We can do even more. We can connect this idea to maps and localization to give the vehicle intent. In this example, we are providing the car with a roadmap. The map is a topological map. Using the same kind of end-to-end approach, the car is able to detect the surrounding environment and follow the road. When it gets to intersections, it's able to match where it is located against what it says in the map, and figure out which road to take.

This is the first ride of an autonomous vehicle in this environment — the vehicle has not been trained explicitly with any of the sensory output that the car sees in this environment. So with these technologies, we are pushing the reasoning ability of robots beyond where they are today, with everything preprogrammed and the car not able to adapt to new situations. By combining techniques in artificial intelligence with robot control, we can achieve new capabilities for autonomous vehicles. Of course, when we do this, we have to keep in mind the error and uncertainty of all the methods.

Now the same kind of approach can be used to enable robots to do many other things. Here you see a robot that is screwing in a light bulb. Now in the future if anybody asks you How many robots does it take to screw in a light bulb?, you know the answer: You need two hands, but the hands have to be soft.

These robot hands that you see here are very compliant. They're made from the same material you saw in the robot fish. This allows the robot to adapt to any shape object without actually knowing what it is that it's grasping, and without needing to know precisely the location of that object. Notice that one hand grasps from the side and one hand grasps from the top. How to grasp is something that the robot has to figure out. Again, this is something we know intuitively, but machines have to get trained; they need reasoning engines to identify the correct approach direction for grasping.

We use a similar approach to what we did with autonomous driving to train the robot to determine the approach direction and the orientation of the hand. In this case, we did not need millions and millions of manually labeled examples for each type of object the robot might need to grasp. What is important for grasping is the aspect ratio of the object. We can train this system with a small set of objects that span different-size boxes. When the system sees a new object, it grasps it in the same way as its enclosing box.

It is interesting to combine machine learning, artificial intelligence and robot control. The end result is a departure from the techniques we have developed over the past 60 years in robotics, which primarily consist of planning robots, where the robots plan in known environments and their reasoning is limited to adaptive reasoning.

Let me move on and talk a little bit about bodies. I want to observe that right now making a new robot is a bit like programming was before the invention of compilers: everything gets done from scratch, and it takes a huge amount of time to do. We would like to accelerate the design and the fabrication of robots. So imagine a future where anybody — for example, Alice — anybody can have a robot. We would like to give Alice the ability to automate any task she wants in her home.

Say Alice wants a robot to play with her cat while she is at work. In this future, Alice would go to a store called 24-hour Robot Manufacturing, where she would get an intuitive interface to figure out a robot design. Within a short amount of time, and for low cost, a robot would be built and now the cat would have a playmate. My research group at MIT is working on this problem. We would like to provide a system with the functional specification of a machine and expect the system to automatically determine a robot design for this functionality.

Say I want a robot to play chess with me. By analyzing the sentence, we could extract the behaviors required by the robot: picking up a part, moving it around without knocking other parts, placing the part. Given these behaviors we can determine mechanisms that implement the behaviors and then synthesize a whole system that includes these mechanisms. Here is the robot that plays chess with you. Now this robot was created in about two hours, and it's made out of paper. It's made by folding, so it's a very inexpensive robot. It was done following a data-driven approach, where we start with a database of many designs, and by segmenting and composing these designs through interactive modeling, we can create new designs that can then be fabricated by printing.

In our system, we are looking at printed fabrication because we want to use rapid 3D-printing techniques. The user gets to reason in three dimensions and the system creating the design operates on a two-dimensional version of that object. All the designs get created hierarchically. Here is a user who's deciding to make this ant-like robot, and the user very intuitively can create the body, can size up the body, can add legs to this body, and then a simulation engine checks the function.

In this case, the robot is not stable, and that gives the user the ability to adjust some of the parameters to make it stable. When the design is finalized, it can be sent to any rapid prototyper like a 3D-printer or a laser cutter. We can take this data-driven approach to design and apply it to the computational substrate of the robot — in other words, to the electronics used in the system. For the ant robot, the user might say, It needs two motors. From this specification, the system is able to automatically figure out the circuit, wiring instruction diagrams, the folding pattern. In other words, everything that is needed, even the programming that is needed to create this robot.

> **"...imagine a world where many objects can be awakened to become active, to become moving and sensing objects..."**

There are other ways in which we can automate the creation of robots. In additive folding, we are given a picture of an object we're trying to create. A compiler slices the object, unfolds it, and creates a long string that can be 3D-printed and folded to create a physical realization of the picture. Folding builds into the object natural hinges and joints that can be actuated to animate it.

Any object in the built environment or in the natural environment can be created from a picture, fabricated by additive folding, and become a roboticized version of your object. So imagine a world where many objects can be awakened to become active, to become moving and sensing objects just like this rendering of the bunny. We might say, What if the Sydney Opera House were a robot? Well, if the Sydney Opera House were a robot, this is what it would be.

[Plays video clip]

Okay, so more opera for later. For now imagine that many of the objects in our surrounding environment could come "alive." Now the objects I have shown you so far were folded by hand. We can even imagine a system where these folded robots make themselves. This is an example of a robot that starts as a piece of plastic and when exposed to heat, the robot grows into a full-blown, three-dimensional machine that can move around. It can do interesting things.

It can run at four body lengths a second, it can avoid obstacles, it can move in different types of environments. And the secret sauce of this robot is the following: this robot is made as a layered structure with structured layers at the top and the bottom, and a layer that is activated by heat in the middle. This activation layer is made from the same material as Shrinky Dinks. By cutting gaps in the top and bottom layers, we can control the angle at which two surfaces lock when creating the 3D structure. The robot has embedded in it a tiny magnet that can be used to control where the robot goes.

Now consider an MRI-like machine that provides an external magnetic field to control the robot. And when we are done with the robot, it can even recycle itself by dissolving in water. Recycling is critical for a world with many different robots.

As we think about more and more machines supporting us with physical and cognitive tasks, we have to think about how we manage their impact on us and on our planet. This little robot that you have just seen can even have superpowers. Imagine taking the robot and using the same folding approach to give this robot different kinds of coats, each coat enabling a different type of movement. This is an exoskeleton that allows the robot to move faster. It allows it to be bigger, it allows it to scoop objects. The robot can remove the exoskeleton. Then the robot could get different exoskeletons to enable rolling, floating on water and gliding. The exoskeletons can be acquired and removed autonomously to enable four different types of locomotion using the same basic substrate for the robot.

We can make these robots at big scales or at small scales. Even at small scales, these robots can be useful. One particular application is in surgery. Think about a future of surgery without incisions, without pain, without the risk of infections.

"As we think about more and more machines supporting us with physical and cognitive tasks, we have to think about how we manage their impact on us and on our planet."

A motivating application is the fact that many people swallow button batteries. Swallowing button batteries is dangerous because within an hour, the acid in the battery causes the battery to get immersed in the tissue, so it punctures. So here's the idea for the solution to removing the battery. You take the mini-origami robots you have seen, you squeeze them in the form of a pill, you put them in ice, and then you swallow them just like you swallow a pill. When the pill-robot gets to the stomach, the ice melts, the robot deploys, and now the robot can be controlled to do useful work. As you can see in this simulated stomach, the robot can be directed using the external magnetic field to pull out the button battery, or later on the robot could serve as a patch for the wound created by the battery, or the robot could deliver medicine very precisely at a location inside the stomach. These robot pills are made out of food — namely sausage casing. They are safe and digestible.

I would like to observe that we have been very constrained in how we think about robot bodies. Most of the robots built over the past 60 years have been either inspired by the human shape, or on wheels. But in fact, nature and the built environment have many other forms. For the future I would like to suggest that we should rethink, What is a robot? Let us rethink the materials from which we make robots and the shapes and the functions of robots. With greater variety of forms, we can achieve a Cambrian explosion of pervasive robots that can support us with many different tasks, because as we observed in the beginning, the form of the robot impacts the kind of tasks it can do.

For the last bit of the talk, I want to say something about interactions, because a world with pervasive robots requires easy interactions between machines and people, and right now, we do not have that. Right now, one has to be an expert in order to use machines. The machines also have to interact with each other. I'm going to skip discussing how machines interact with each other, and I'm going to focus on intuitive interactions with people. I want to start by showing you a project that's very dear to me, and this project was catalyzed from a conversation with Andrea Bocelli, who is blind. To enable safe navigation for the blind, the hardware and software that enable autonomous cars to move safely in the world can be mapped onto wearables. The laser scanners are added to a belt. The belt has a belt buckle that can talk to the blind person. There is a camera. It's really nice, fashionable-looking — worthy of Italian fashion. This laser-camera system can perceive and understand the environment, first of all for safe navigation but for much more than that. This system can describe a fabulous window display, say, Hey, there's an obstacle behind you, or that a friend is walking by.

Here is a blind user walking along a narrow hallway, up and down the stairs, and even outside, avoiding obstacles without bumping into anything. With intuitive feedback, the system directs the blind user to a bench, where the user could sit and feed the ducks. This type of capability could bring tremendous quality of life improvement to the blind and visually impaired community. We have the ability to use our navigation technologies coupled with intelligent interfaces to deliver safe navigation to visually impaired people.

"...we have been very constrained in how we think about robot bodies. Most of the robots built over the past 60 years have been either humanoid, inspired by the human shape, or they have been in arm form or maybe on wheels. But in fact, nature has so many different forms, and the built environment has so many different forms...I would like to suggest that we should rethink, What is a robot?"

Ultimately, we would like to have machines adapt to us rather than the other way around. While today we use natural language to program machines, a more advanced question is, Can we connect to the machine directly through our brain waves?

Is it possible to monitor what goes on in our heads? Is it possible to monitor brain activity and extract something useful from that that can be used as a control signal to interact with machines?

The answer is that in general, no. But it turns out that even with a sparse sensor like the EEG cap, there is one signal that can be detected reliably. This is a signal we all make, no matter what language we speak in. It is the signal our brain makes when we notice that something is wrong. In other words, the "you-are-wrong" signal or the error-related potential. This signal has a unique signature, it's localized, and it can be detected, and it can be used to enable intuitive interactions between people and machines. Here is an example where a user is watching a robot. This robot is tasked to sort paint cans in a bin labeled "paint" and wire spools in a bin labeled "wire." When the robot places the object correctly, the user does not generate any error-related potential, but when the robot makes a mistake, the error-related potential is detected. That signal can be detected in 100 milliseconds, and then it can be used to change the behavior of the robot. Here, the robot aims to place the wire spool into the paint bin, which is wrong. In 100 milliseconds, the robot was able to correct its mistakes using the signal from the human. I believe there is potential for a future where, instead of creating machines that we have to adapt to, we have machines that will adapt to us, that will work in sync with our needs, in sync with our desires.

"I believe there is potential for a future where, instead of creating machines that we have to adapt to, we have machines that will adapt to us, that will work in sync with our needs, in sync with our desires."

So we talked about the science and engineering of autonomy, we looked at brains, we looked at making a robot more capable of figuring things out, we looked at making robots faster using computational design and fabrication, we looked at different kinds of interaction for the future between humans and robots. Where might that take us? Well, imagine a future where you wake up, enabled by your personal assistant that figures out the optimal time. The assistant organizes everything you need for the day while you sleep. It even figures out the appropriate outfit for the day, given your schedule and weather. As you walk to work by storefronts, imagine new technologies that display your picture with the latest fashion in the store windows.

When you walk inside the store, a 3D-scanner scans your body and creates bespoke clothing and footwear for you. Today, we even have computational fibers that in the future can be used to create programmable clothes. When two friends meet, they could reprogram their outfits made of computational fabrics to match each other's sweaters. At work, there are intelligent rooms that monitor people closely and adapt the ambient environment and the types of interactions and the particular discussion topic.

Physically present people and virtually present people are able to manipulate cooperatively a prototype for the new flying car. This flying car is connected with the city infrastructure. Alice's mother here gets a message to pick up some plants from the nursery, which is having a sale. Well, the car could immediately figure out a small change in its travel plan to go to the sale. Back at home, as Alice's little sister takes the first steps riding a bicycle, the training wheels adjust just in time, robots deliver goods, garbage cans take themselves out, robots can help with planting, all of this to enable Alice's little sister and her grandmother to spend more quality time together.

After a good day, it is time for a bedtime story. The technologies of the future will enable children to enter the story and interact — in this case with dragons.

The recent advances are really shaping our future with machines. Imagine a future where if you think of an object, you can actually make it. Imagine a future where any kid has the ability to make their own version of the underwater sea observatory. All of us, in some sense, could use the new ideas of designing, fabricating and using robots to create new machines that will help us save lives, that improve our quality of life, that will entertain us, that will take us to places we cannot go to. I believe that this future with pervasive robots has in it a lot of extraordinary potential, and I believe that this future is not that far off. We need all the help from the young students all over the world to help us push the technology forward to get to this world of pervasive machines helping us with cognitive and physical work.

Thank you very much.