

PATRICK WINSTON: Well, don't stop.

Shoot.

I guess we've got to stop.

I will soon go into withdrawal symptoms that will last about six weeks.

But, on the other hand, we all are beginning to develop a sort of tired and desperate look.

And perhaps it's a good thing to get the semester behind us and go into solstice hibernation.

Anyway, there's a lot to do today.

I want to wrap up a couple things, talk about what's next, and maybe get into some big issues, perhaps a little Genesis demo, that sort of thing.

So here's what we're going to do first.

Last time, I talked about this whole idea of structure discovery.

And really, the whole reason I cracked and started talking about basic methods is because of the potential utility of taking that idea one step further and finding structure in situations where you might not otherwise find it.

It's still an open question about whether that's best way to think about it.

But here it goes.

Imagine you've got a couple of stories.

And these circles represent the events in the story.

And now what you'd like to get out of these stories is some kind of finite state graph that describes the collection of stories.

So you might discover, for example, that these two events are quite similar.

And these two events are quite similar.

So you might use that as a basis for speculating that maybe a more compact way of representing the stuff in the story would look like this.

Where this one-- let's see.

This one goes with this one, this one goes with this one and there's a possibility of another state in between.

So that's the notion of Bayesian story merging.

Now I'd like to show you a little bit more convincing demonstration of that.

Here's how it goes.

So there are the two stories.

This is just a classroom demonstration, no big deal.

But you can see there's a sort of parallel structure in there.

So this is the work of a graduate student, Mark Finlayson, who processed those stories to produce those kinds of events and those kinds of events that get assembled into two story graphs.

And the question is, is the most probable way of explaining that corpus of stories?

And of course, the answer is no.

If you merge some things like chase and stop, then you get a simpler graph, one that is more probable in the same sense that we discussed last time.

Then you can merge run and flee because they're similar kinds of events.

And finally, you've got think and decide.

Boom.

There is your story graph.

And this is the same idea taken several levels higher that produce the capacity to discover, in these two stories, the concept of revenge, as promised at the beginning of our last discussion.

So sometimes the Bayesian stuff is the right thing to do, especially if you don't know anything.

But sometimes you do know stuff.

And when you do know stuff it's possible that you can do something very much more efficient.

This sort of thing takes clouds of computers to process.

But we learned a lot of stuff in the course of our development that we don't use a cloud of computers to figure out.

We learned how to associate the gestures of our mouth with the sounds that we make, for example.

So I want to spend a minute or two talking about some work that someday will be the subject of a couple of lectures, I think.

But it's the question of how to use multiple modalities and correspondences between them to sort out both of the contributing modalities.

That sounds contradictory.

Let me show you an example.

This is the example, a zebra finch.

And it's showing you the result of a program written by Michael Coen, now a professor at the University of Wisconsin.

So the male zebra finch learns to sing a nice mating song from its daddy.

And this is what one such zebra finch sounds like.

[BIRD SONG PLAYING] Nice, don't you think?

And here's what was learned by a program that uses no probabilistic stuff at all, but rather the notion of cross-modal coupling.

[BIRD SONG PLAYING] Can you tell the difference?

It's not known if this particular song turns on the female zebra finch.

But to the untrained human ear, they sure sound a whole lot alike.

So how does that work?

Here's how that works.

Well, I'm not going to show you how that works.

What I'm going to show you is how the classroom example works, the first chapter example in Coen's Ph.D. Thesis.

Here's what happens.

When we talk, we produce a Fourier transform that moves along with our speech.

And if we say a vowel like aah, you get a fairly constant Fourier spectrum.

And you can say, well, where are the peaks in that Fourier spectrum?

And how do they correspond to the appearance of my mouth when I say the vowel?

So here's how that works.

So here's the Fourier spectrum of a particular vowel.

And when you smooth that, those peaks are called formants.

And so we're just going to keep track of the first and second formant.

But when I say those things, I also can form an ellipse around my mouth when I say them.

And when I form an ellipse around my mouth when I say them, that gives me this second modality.

So the question is, is there a way of associating the gestures that produce the sound with the sound itself?

Well, there's the human data conveniently provided by a variety of sources, including Michael Coen's wife who produced the lip contour data on the right.

So that's all marked up and color coded according to the particular vowels in English.

I guess there are ten of them.

So we humans all learn that.

But guess what?

We don't learn it from this.

Because we don't get to work with any marked up data.

We learn it from that.

Somehow we're exposed to the natural world, and we dig the vowel sounds out.

It's fantastic how we do that.

But we do have cross modal coupling data and maybe that's got something to do with it.

So here is a particular cluster of sounds.

And what I want to know is, can I merge any of these two clusters to form a bigger cluster with a corresponding meaning?

So what I can do is, I can say, well, I can watch these.

I know what the lip form is when a particular sound is made.

So I have these correspondences.

So maybe there are four of those, like so.

And maybe this same guy projects a couple of tones into that.

And the question is, can this guy be combined with any of these guys?

And the answer is yes.

If they're close together on one side, maybe that suggests you ought to cluster them on the other side.

But there's a question about what close means.

So let's suppose that we also look at how these guys, these other two guys, project.

And suppose this guy projects twice up here and once over here.

And this guy down here just projects like crazy into that guy.

Which are closer?

These two or these two?

Well, my diagram's getting a little closer.

But if you paid attention when I was drawing it, you would see that this guy projects in proportion to that guy.

So if we look at the-- if we take each of these projections as the components of a vector, then those two vectors are in the same direction.

And the cosine between them is zero.

So these are the two that are closest together from that sort of perspective of that kind of metric.

And those guys are the ones who get combined.

Would you like to see a demonstration?

Yeah.

OK, here's a demonstration based on Coen's work.

So here we have two sides.

We could think of one side as being vowel sounds and the other side as being lip contours or something.

But you don't see anything in the diagram so far about how these things ought to be sorted out into groups.

So if I just take one step, why, it discovers that those two guys had the same projection pattern as each other.

So if I take another step and do the same thing on the other side, and now in the third step, the two areas that were formerly combined now form a super area.

And they're seen to project in the same way as the blue area.

So using this kind of projection idea, I can gradually build up an understanding of how those regions go together.

And I discover, in this contrived example, that there's a vertical arrangement on the right side that corresponds to a horizontal arrangement on the left side.

Now, you say to me, I'd like to see something a little bit more like the lip contour data.

I'm just stepping through here until I get something I kind of like.

Oh, that sounds good.

That seems good.

So there's a correspondence here.

This is all made up data, Gaussians of various shapes and orientations.

Let's see what happens when I run the clustering algorithm on that.

Something definite was learned at every step.

We find the correspondence between the pink region on the right and the pink region on the left.

In some cases, where the regions are rather blurred together, the other side is the one that helps the system figure out how things are organized.

So I cite this as an example of something I think is very important.

Number one, it's possible to discover regularity without being obsessively concerned with Bayesian probability.

And also, that there's very likely a whole lot of this going on in human intelligence.

When we emerge and begin to examine and explore the world around us, we're presented with a lot of unlabeled data that we've got to make sense of.

And I believe that this kind of cross modal coupling idea is very likely to be bound up in our understanding of that world that's presented to us.

It's fast, it's direct.

It doesn't take thousands of data points.

It just happens.

And it happens effortlessly.

And if this isn't built in-- if this isn't determined to be built in, you can come back to MIT in 15 years and put me and jail.

Because I think this is really the way it works.

So there it is.

There's a couple of things to have wrapped up.

And now the next thing I want to do for the rest of our last time together in this format is talk to you about a variety

of things.

And I'll depart from my usual practice and move to some slides.

So Dave, could we have the center screen, please?

So first, a brief review of where we've been and where we've come.

I think in the very first class, I talked about what artificial intelligence was.

And I talked about how you could view it from either an engineering perspective or a scientific perspective.

I'm on the scientific perspective side.

And I think, nothing against applications, but I think we'll be able to make much more sophisticated and wonderful applications if we have not only the engineering perspective about building stuff but also the scientific perspective about understanding the stuff to begin with.

So both perspectives are important.

And in this case you can see that they all involve representations, methods, and architectures.

Dave, I've changed my mind.

Could you also give me the side screen so I can see it too?

So, that's that.

What's next?

The business perspective, which we talked about on Thanksgiving.

The important idea being that the knee-jerk expectation that the commercial value of something is in replacing people is something that is not sensible in the first instance, and demonstrated to be unlikely and untrue in the second instance.

The thing that turns people on from the point of view of applications is not replacing people but making new revenue, making new capability.

And that at once licenses you to not have something done exclusively by a computer but something that can be done in partnership with a person.



So all the important applications of artificial intelligence involve people and computers working in tandem, with each doing what they do best-- not with replacing people.

So that's that.

Here's what AI does that makes AI different from the rest of the fields that attempt to contribute to an understanding of intelligence.

Now, we have the benefit of having a language for procedures.

We have all the metaphors that we are the custodians of in consequences of knowing about programming.

We have the metaphor of garbage collection.

We can talk about all sorts of things with programming metaphors that are unavailable to people in other fields that are interested in psychology.

We have a way to make models because we can write programs.

And when we write programs, there's no question of sweeping things under the rug.

We have to work out the details.

And once we've done all that, then we have opportunities to experiment that are beyond the ability to experiment in most other fields.

Oh, magnificent experiments are done these days in developmental psychology and all the rest of-- all the other branches of psychology, including MRI studies that probe into your brain and see how it's consuming sugar.

But it's very difficult to ablate or take away some piece of your knowledge and see how you work without it.

I can take the line-drawing program that we talked about and say, how will it do if it doesn't know anything about fork junctions?

And we can determine an answer.

But I can't reach into Sebastian's head here with a surgical procedure and take out his knowledge of fork junctions.

It just can't be done.

And finally, another reason why we're different is because we can put upper bounds on how much knowledge is needed in order to perform a certain kind of task.

These days, with bulldozer computing, the question most often asked is, how can you get billions of the things off the web and use them.

We-- I especially-- sometimes ask the opposite question, which is how little knowledge can you have and still understand a story?

That's what's interesting to me.

So there's a methodological slide that talks to the question of how you do artificial intelligence in particular and, I suppose, science in general, engineering in general.

There's a great tendency in this field to fall in love with particular methods.

And we've had people who've devoted entire careers to neural nets, genetic algorithms, Bayesian probability.

And that's mechanism envy.

And a better way, in my judgment, is to say, what is the problem?

Scientific method-- what's the problem?

And then bring the right machinery to bear on the problem, rather than looking for things to do with a particular kind of machinery.

So this is the methodology that first articulated in a forceful way, by David Marr.

You want to start with the competence you're trying to understand, then bring a representation to bear on it, a representation that exposes the constraints and regularities.

Because without those, you can't make models.

And without those models you can't understand it, explain it, predict it, or control it.

So it seems to make sense from a kind of MIT, model-centered point of view.

And only when you've got all that straight do you start working on your methods and implement an experiment and then go around that loop.

So that's all I want to say by way of review, I suppose.

I want to take a minute or two and just remind you of what's on the final.

And there's nothing to remind because you all know what's on the final already.

We'll have four sections corresponding with four exams.

Then we'll have a fifth and final question that will be everything else.

All that stuff you slept through will be featured there, as well a little problem on Bayesian inference.

We rearranged the subject, mostly so I could write the demonstrations.

So that the Bayesian stuff didn't come before the fourth quiz.

Therefore the Bayesian stuff that you see on those previous quizzes is likely to be harder than the stuff that we'll ask on the final, because you haven't had as much experience with it as people did last year.

I've got a few icons on there to remind me to tell you a few things.

As always, open everything except for computers.

You can wear a costume.

You can do anything you like as long as it doesn't disturb your neighbor, within reason.

Well, I guess if it doesn't disturb neighbor, it is within reason.

So maybe that's all I need to say.

I'm not sure where we're going to be.

But it's certainly the case that, historically, there are no visible clocks.

So, we soon run out of all of our cellphones, wrist watches, and other time pieces as we hand them out.

So it pays to remember to bring some kind of timepiece, because we won't be able to convey the time very well.

And finally, I see a little calculator there.

I don't recall any exam where you actually needed a calculator.

But it's sort of a security blanket to have one.

People sometimes see a problem and say, oh my God, what am I going to do?

I left my calculator at home.

So as to avoid that anxiety, you might want to bring one even though you won't need it.

So that's the final.

I'm sure there are no questions.

Are there?

It's obvious.

Everybody will do well.

Two shots, that whole thing.

Now, what to do next?

Suppose this subject has turned you on.

There are a variety of things that you should be thinking about doing next semester.

And I wanted to review just a few of those.

One of them is Marvin Minsky's subject, Society of Mind.

It's very different from this class.

There are no prepared lectures.

Marvin doesn't rehearse.

He doesn't think about what he's going to say in advance.

It's like this except it's just a conversation with Marvin.

So many people find themselves bored stiff for two lectures out of three.

But then in the third lecture, Marvin will say something that you'll think about for a year or for the rest of your life.

That's what happens to me.

I'm bored stiff two out three times.

And then the third lecture he says something, and I think about it for at least a year and maybe permanently.

So it's an opportunity to see one of MIT's true geniuses think out loud.

So it's an experience that you don't want to miss because that's what you come here for, is to see the geniuses think out loud.

Speaking of geniuses, then there's Bob Berwick.

And he heroically is doing two subjects in the spring.

Both of which I'd take if I could.

One is his subject on Language Understanding.

And the reason I'd take that is because I believe that language is at the center of any explanation of our intelligence.

So that's the subject I would be, I suppose, most inclined to take if I were you.

Well, maybe Minsky's.

It's hard to say.

And incidentally, very heroically, Bob is also teaching a course on how evolution works, how it really works-- in so far as we know how it really works-- as well.

So both of those will be offered in the spring.

I don't know how he does it.

I don't know how he does two all at the same time.

Of course there are lots of places where you can go to school, and the faculty will be teaching five courses at the same time.

I just think they're crazy or something.

I don't know how that works.

Gerry Sussman will be teaching his Large Scale Symbolic System subject.

That's sometimes-- oh, I forgot what he wasn't able to call it, something that wasn't politically correct about programming for people who really, really like to program.

It's a splendid course on how to build really big systems.

And we use the ideas in that subject in our research system, because it's the only way-- understanding how that works is the only way that you can build systems that are too big to be built.

I may say a word about that a little later.

So those are my favorite three/four picks.

But there's lots of other stuff, too many things to cover-- the media lab, [INAUDIBLE] psychology.

There's tons of stuff out there.

And I would only mention that courses that have those three names on them are bound to be good.

These are colleagues that I think I have a important perspective-- not necessarily one I agree with, but an important perspective that you should understand-- Richards, Tenenbaum, and Sinha.

And now we come to my spring course, the Human Intelligence Enterprise.

It's 6.XXX not because there's anything pornographic about it but because for a long time, I couldn't remember it's number.

So I developed a habit of referring to it as 6.XXX and it seems to have stuck.

Here's what that's about.

Yeah, that might be interesting.

It's taught like a humanities course, though.

No lectures, I just talk.

And all the TAs are veterans of that class so if you want to know if you should do it, you have several resources.

You can talk to them.

Or you could look at the sorts of things we talk about.

Here are some things we talk about by way of packaging.

Yeah, I can hear a little tittering there because people have discovered the last element.

Some people take the whole subject because they want to be present for that unit.

And we talk about all those kinds of things.

And we look to see what the common elements are in all those kinds of packaging problems of the sort that you will face over and over again when you become an adult, no matter what you do.

If you become a business person, an entrepreneur, a military officer, a scientist, or an engineer, that packaging stuff will often make the difference between whether you succeed or don't.

And then, that's the second way you can figure out whether you want to take the subject.

The content, the TAs, and then of course you can always appeal to the Underground Guide.

And that's why it's very rare for someone that takes 6.XXX who hasn't been at this final lecture because they read the Underground Guide.

Here is an element that appeared in the Underground Guide a few years back.

There are no exams.

But there is a tradition of hacking the Underground Guide.

So this is another example of something that appeared.

So it all came about because early in the teaching of 6.XXX, I was whining to the students about the fact that I've been at MIT for a long time, since I was freshman.

And I still have yet to have any person I report to say anything about my teaching-- good, bad, indifferent.

Nothing.

Not a word.

So the students decided that it would be interesting to see if they could say something sufficiently outrageous to force a conversation between the department chairman and me.

And so far they've been totally unsuccessful.

And I've tried everything.

Winston shows up late if he shows up at all.

Good instructor but constantly sipping from a brown paper bag.

All kinds of stuff.

But there it is.

It's a lot of fun.

It's a little oversubscribed so we have to have a lottery and there's about 50% chance and so on and so forth.

But many of you will find it a good thing to do.

Oh, yeah.

And now I also want to remind myself that there is a IP event that's become kind of MIT tradition.

It's the How to Speak lecture that I give.

This year it'll be on January 28 in 6120.

6120 holds about 120 people and about 250 show up.

So if you want to go to that lecture you should show up 15 minutes early.

That's a little secret just between me and 6034 students.

It's about packaging, too.

It's a one-lecture version of 6.XXX.

But it's very nonlinear because one thing that you pick up from that one hour may make the difference between you getting the job and some other slug getting the job.

So it's one of those sorts of things that can make a big difference in a short period of time.



You may sleep through 50 minutes of the 55 minutes in that lecture and stay awake for that one magical five minutes when you learn something about when to tell a joke or how to open a lecture or how to conclude one or a job talk or a sales presentation or anything.

And that will make it worthwhile for you.

And then of course there's the possibility of your-- anybody who's doing [? UROP ?] with me is likely to be interested in what I've recently come to call a strong story hypothesis, something that we've talked about from time to time.

That's what makes us humans, and that's not what they are.

They're orangutans or chimpanzees, even though the DNA that we share is-- it goes up and down.

At one point it was 96%, then it went to 98%.

Now I think it's back down to 97%.

But whatever we are, it's not because of a huge, massive difference in DNA between us and our cousins.

So in my group we build a Genesis system, modestly called.

And that's what it looks like.

And it has in it all the sorts of things that we've talked about from time to time in 6034.

And it's about to move into areas that are especially interesting, like can you detect the onset of a possible disaster before it happens and intervene?

Can you retrieve presses based on higher-level concepts?

Things of that sort.

Would you like to see a demonstration?

OK.

So you've see in a little bit of this before.

In fact, I'm not even sure what exactly I've already shown you.

Let me just get over there to see what goes on here.

What I'm going to do right now is I'm just going to read about Macbeth.

A short precis of the plot.

Not the whole thing, of course.

Just a few sentences about the plot.

Right now what it's doing is absorbing the English and translating into a sort of internal language.

A sort of universal, internal language that's all about trajectories and transitions and social relationships and all that.

So it's being read there by two different persona.

Each of those personas has a different educational background, you might say.

They might represent different cultures.

So eventually they build up graphs like that.

And everything in white is stuff that's explicit in the story.

And everything that's in grey is stuff that's been inferred by the system.

So there are several layers of understanding.

One is what's there explicitly.

And the other thing that's there is stuff that's readily inferred.

And because these persona have different educational backgrounds, you might say they see the killing of Macbeth at the end of the play in a different light from one another.

One sees it as an act of insane violence and the other sees it as a consequence of a revenge.

So once you've got that capability, you can do all sorts of things.

For example, you can ask questions.

So let me arrange it to ask-- by the way, this is a live demonstration.

I'm thrilled to pieces that it actually works so far.

Let's see.

Why did Macbeth kill Duncan?

You all know why, right?

Yeah, you're right.

He didn't actually do that.

On a common sense level, neither Dr. Jeckll nor Mr.

Hyde have an opinion.

On a reflective level neither Dr. Jeckll more Mr. Hyde have an opinion.

That's because it didn't happen.

So we call these two persona Doctor Jeckll and Mr. Hyde.

And you're ready to complain right away about the spelling of Jeckll, aren't you?

Well, that's because with this spelling the speech generator makes it sound a little bit more like when we say Jekyll.

But what really happened is that Macduff killed Macbeth.

On a common sense level, it looks like Dr. Jeckll thinks Macduff kills Macbeth because Macduff is insane.

It looks like Mr. Hyde thinks Macduff kills Macbeth because Macbeth angers Macduff.

On a reflective level, it looks like Dr. Jeckll thinks Macduff kills Macbeth as part of an act of insane violence.

It looks like Mr. Hyde thinks Macduff kills Macbeth as part of acts of mistake, Pyrrhic victory, and revenge.

Isn't that cool?

I bet you'd get an A if you had said that in eighth grade.

But once you've got this ability to understand the story from multiple points of view, you begin to think of all kinds

of wonderful things you can do.

For example, you can have Dr. Jeckll negotiate with Mr.

Hyde, because Dr. Jeckll will be able to understand Mr.

Hyde's point of view and demonstrate to Mr. Hyde that he thinks that point of view is legitimate.

Or, Dr. Jeckll can teach Mr. Hyde the subject matter of a new domain.

Or Dr. Jeckll can watch what's happening in Mr. Hyde's mind and avert disaster before it happens.

So let me just show you another situation here.

I want to turn on the onset detector and read another little snippet.

This one is about-- what should I do?

We'll do the Russia and Estonia cyber war.

It's reading background knowledge right now.

But pretty soon, in the upper left-hand corner, as it begins to read the story, you'll see it spotting the onset of potential revenge operations or potential Pyrrhic victories, and show their foundations begin to emerge and giving the system an opportunity to intervene.

So there you can see all the things that it thinks might happen.

Not all of them do happen.

But some of them do.

And you'll note, incidentally, that this is another case of Dr. Jeckll and Mr. Hyde having different cultural perspectives.

One's an ally of Russia and one's an ally of Estonia.

One sees it as unwarranted revenge and the other sees it as teaching a lesson.

So, I don't know.

What else have we got here?

Oh yeah, president recall.

A long time ago, we talked about doing information retrieval based on vectors of keyword counts.

That's cool but not this cool.

This is doing it on vectors of concepts that appear in the stories, such as revenge, even though the word revenge doesn't appear anywhere.

So because we're able to understand the story on multiple levels, we can use those higher levels that don't involve the words in the story at all to drive the retrieval process.

So all that is a consequence of a variety of things, one of which is the specialists that translate the English into an internal language.

And it's also, incidentally-- I mentioned it a little before-- it's also a consequence of our use of Gerry Sussman's propagator architecture.

So a student comes into our group and says he wants to do something.

And we say OK, here's how the system is organized.

It's like a bunch of boxes that are wired together.

So you get a box and we'll tell you what the inputs are going to look like.

And we'll tell you what we want on the outputs.

And if you don't like the inputs, just ignore them.

And if we don't like your outputs, we'll just ignore that.

So nobody can screw up anything because they have a very circumscribed piece of the system to work with.

So I can say, for example, the president wanted Iraq to move toward democracy.

And bingo.

That starts a propagation through that network.

All this would be unconvincing, in my view, if it weren't eventually connected to perception.

Because if it's not eventually connected with perception, it's yet another system that demonstrates how smart it can seem to be without actually knowing anything.

So another half of what we do is an early stage attempt to connect the language stuff with things that are going on in the world.

So we say, imagine a jumping action.

And there is a jumping action that's part of a test sweep developed by the Defense Research Projects Agency to drive what they call the Mind's Eye program, which was developed largely as a consequence of work done here at MIT, focused on the idea that if we're going to understand the nature of intelligence, we have to understand how language is coupled into our perceptual systems and how those perceptual systems can answer questions posed to them by the language system.

That's a little demo of the Genesis system.

Here are the issues that we're trying to explore.

Nothing too serious, just the nature of what is extraordinarily fundamental to any explanation of human thinking.

Now, all of this might turn you on.

And you say to me, well, you're sick and tired of MIT.

You'd like to go somewhere else for graduate school.

So now that I've demonstrated what we do here, one of the many things we do here, I'll talk a little bit about other places you can go.

This is a sort of MIT-centric view of the world.

It represents all of the places you could go when I was a kid.

But while I've got this particular diagram on here, I just-- sort of testing my MIT arrogance-- I remember a story often told by my colleague Peter Szolovits.

He says that when he came to a job interview from Caltech to MIT, he was sitting here for three days and nobody spoke to him.

So eventually he said, I've got to do something.

He walked up to a graduate student and said, hi my name is Peter Szolovits.

I'm from Caltech.

And the graduate student said, Caltech sucks, and walked away.

Anyway, we've populated all these places now that you see here, and more.

This is just a list that I scratched up this morning.

I'm sure I've forgotten many that have equal right to be on this list.

But in the end, which one you go to depends on who you want to apprentice yourself to.

Because a graduate school is an apprenticeship.

And that means if you go to a place with just one person, it's OK if that's the person you want to apprentice yourself to.

Each of these places has a different focus because they have different people.

So you need to find out if there's somebody at any of these places.

It doesn't matter if it's AI or some other field.

Theoretical physic-- you've got to find out if there's somebody at that place you want to apprentice yourself to.

So those site visits are really important.

And I would like to also stress that when you make your application to graduate school, it's very different from applying to undergraduate school.

Because they don't care whether their school is good for you at all.

They only care about one thing-- whether you're good for their school.

So don't get confused and talk about how it's a wonderful fit for you.

Because what they're interested in is whether you're going to contribute to their research program.

Oh, I should say that if you're applying to artificial intelligence that means you don't say, I'm interested in all aspects of thinking.

You need to be focused.

There's another reason why you don't say that you're interested in all aspects of thinking and that is the defect theory of AI career selection.

It seems to be the case, strange though it may seem, that people in artificial intelligence often specialize their research on the things that they don't do very well themselves.

So people who study language, with the exception of Bob Berwick, often have trouble getting out a coherent sentence.

And people who do hand-eye coordination are the sorts who spill their coffee.

So don't say you want to study all thinking because-- The most extreme case of this, though, is-- if you don't mind, I'll tell you a story about an extreme case in this.

We had a visitor from Japan in the old artificial intelligence lab many years ago.

He came for a year.

Let's call him Yoshiaki, just to pick a name.

Yoshiaki spent a year at the artificial intelligence lab, and he left his wife in Japan.

And the reason was, she was pregnant.

And at that time, you could not get a visa to the United States unless you had a smallpox vaccination.

And because she was pregnant, she didn't want to get a smallpox vaccination because there's a small danger to the fetus if you get a smallpox vaccination while you're pregnant.

So she stayed back there.

So Yoshiaki, let us call him-- it was a day before he was to get on the airplane to go home.

I walked into his office and his desk was covered with pictures of his wife.

By the way, Yoshiaki, I should tell you, is a computer vision guy, interested in object recognition.

So you might suspect he has some problem.

So he's looking at these pictures.



I thought, oh my God, this is a tender moment.

He's anticipating his return to Japan and reunion with his wife.

So I muttered something to that effect.

And then he looked at me like I was the king of the fools.

And he said, it's not a question of tenderness.

I'm afraid I won't recognize her at the Tokyo airport.

So I said, Yoshiaki.

How can this be?

You study computer vision.

You study object recognition.

This is your wife.

How can you think you wouldn't recognize her at the Tokyo airport?

And then he looks at me, and-- God is my witness-- he says, they all look alike.

Well now as we come close to the end, what are the big questions?

Is it useful?

Of course it's useful.

It's part of the toolkit, now, of everybody who claims to be a computer scientist.

What are the powerful ideas and these things?

Well, here's the most powerful, powerful idea is the idea of powerful idea.

And here are a few of my favorites.

No surprises there.

That's just Winston's picks.

But there's one more I would like to add.

And that is all great ideas are simple.

A lot of times we at MIT confuse value with complexity.

And many of the things that were the simplest in this subject are actually the most powerful.

So be careful about confusing simplicity with triviality and thinking that something can't be important unless it's complicated and deeply mathematical.

It's usually the intuition that's powerful, and the mathematics is the [INAUDIBLE] element.

Sometimes people argue that real intelligence is possible.

One of the most common arguments is, well what if we had a room?

And you're in the room and you're asked to translate some Chinese documents.

You've got a bunch of books.

And in the end you could do the translation.

But you cannot be said to understand Chinese.

This is the argument of Berkeley philosopher named Sorel.

So the trouble is, it's also true-- well, the argument is the books aren't intelligent.

They're just ink on a page.

And the person is just a computer, just a processor.

It doesn't actually know anything.

So since it can't be in either the person or the books, it can't be.

And that just forgets that there's a magic that comes about when a running program, when a process, executes in time over knowledge that it continually contributes to.

So the reductionist arguments are among the many that have been ineffectually posed to argue that artificial

intelligence is impossible.

But that bears longer discussion.

Let me just bring up the biggest issue in my mind, which is, it's not the question of whether we humans are too smart to have our intelligence duplicated or excelled in a computer.

It's a question whether we're smart enough to pull it off.

I once had a pet raccoon.

Now, it's illegal to have a pet raccoon.

But this one was an orphan.

Its mother had been hit by a car or something.

A friend of mine brought the raccoon to me knowing I kind of like animals.

And I have to say, I kept the raccoon for a year.

At that point, she wanted to go out and be on her own.

So I had this raccoon.

And this raccoon is smarter than any dog I've ever had.

Within a day, she learned how to pry the refrigerator door open.

So I spent that whole year taping the refrigerator door shut every time.

And then, we'd play jokes on each other.

She wouldn't eat hot dogs.

And I wanted her to eat hot dogs desperately because they're cheap and easy to serve.

All she would eat was cooked chicken wings, wouldn't eat hot dogs.

So one day I said, well, I'm going to play a trick on her.

I took a chicken bone.

I stuck it in the middle of a hot dog and put it in a garbage can.

And she went for it.

Her genes took over and she went for it.

And she was happy with hot dogs ever after.

She wouldn't let me read.

She would crawl up underneath the book and interfere and make me-- she always wanted to suck on my thumb, which turned blue eventually.

You'd be amazed at how much a raccoon can suck.

It's just extremely powerful.

The best parts were when she would go bike riding with me.

I put on a heavy sweater because they have a pretty good grip.

I'd put on a heavy sweater and she'd kind of mount herself on my back and look out over my shoulder, stopped traffic for miles around.

So she was really smart.

But the interesting thing is that at no point did I ever presume that that raccoon was smart enough to build a machine that was as smart as a raccoon.

So when we think that we can, it involves a certain element of hubris that may or may not be justified.

Well, there it is.

Just a couple more things to do.

One of which is, you should understand that Kendra and Kenny and Yuan and Martin and Gleb are doing a lot of stuff that's outside their job description.

All of these quiz review deals that they've arranged are not in their job description.

I didn't ask them to do it.

That's all just plain old professionalism.

So they've been wonderful to work with and I'd just like to-- [APPLAUSE] --offer them a round of applause.

And of course, Bob and Randy and Mark have done fabulous stuff as well.

And we of the staff, the TAs, Mark, Bob, and Randy, have nothing else to do except wish you good hunting on the final and a good long winter solstice hibernation period after that.

And that is the end of the story.

And we hope you live happily ever after.