[SQUEAKING]

[RUSTLING]

[CLICKING]

MICHAEL CUTHBERT: OK, so today we're just going to go through some topics. And stop me when you are. I want to-- the first thing I'll do is just talk a little bit about some of the things in my research. And some of them are connected with a topic that's lingering in the back of this class that but that's not really explicit. And that is the second adjacent field of statistical and pure math, computational musicology research that is not directly computational. So we'll do a little bit with that.

You'll find on the Canvas site there's a very cool-- I think it's like 9-page-- it's really short-- it has lots of pictures-- article that I'd like you to take some time to read by Leon Harkleroad, who wrote a whole book on mathematics and music, which is very cool.

But I find the most cool thing about that is the final chapter, how not to mix math and music, and the ways people have mixed math and music and got things really wrong. So I might bring that up a little bit more. So we'll do a little bit with some statistical and math things that relate to my own research, too.

So I'm trained as a-- actually, let me turn off this light somehow. Yep. Is that good? I'm trained as a medieval music historian, so I really, really got into medieval music. You'll hear little clips of it here and there. And that was really what I wanted to do when I programmed on the side to make some spending money and things like that.

When I got into medieval music, I thought I'd be looking at things that are really beautiful like this all the time. These are people playing a little crank organ, hurdy-gurdy type thing, and stuff. And there are these beautiful, beautiful music manuscripts. This is from the early 14th century called the *Novel of Fauvel,* the *Horse Named Fauvel,* the *Roman de Fauvel.* And they're all really, really gorgeous.

Or there's the original "Carmina Burana" that Carl Orff based his music off of. This is what music notation once looked like. And then gorgeous books in the shape of a heart.

So I started focusing on the 14th century in Italy, where we have this one really amazing book. This is called the *Squarcialupi Codex.* Has these beautiful pictures of all the musicians who are in there. And it's a great-- I think it's valued at around $35 million or something like that.

So I thought I'd be working on this. Well, it turns out people have been doing medieval music research for, I don't know, hundreds of years. And guess what? They've all studied the really pretty books. So if you wanted to go into this in around 1998, 2000, whenever I really got it started, about 25 years ago, here's what's left for us to study.

Little things like that. You can kind of see the music here, and stuff like that. Or this is one of the ones I've been looking at. Some of these, worms eat them, ravages of time, or other little indignities. The music is wrapped up to keep people's sewage payment records. You have to pay somebody to clean out your chamber pots, and you want to make sure that somebody paid their bill. And so you wrap it in a piece of worthless music, and you do that.

Well, when you're wrapping music on top of other documents, sometimes the document's smaller than the music. So no problem. You trim the music, and then it fits there. And some people-- there's nothing important on the left side of a score, right? Just clefs, time signatures, key signatures, nothing like that. And they used a lot of clefs back then.

And then sometimes you wanted things to look really nice. So you would wipe off, wash off, or scrape off all the music that was there so that your cover looked a little nicer. And then once you've scraped everything off, it's all parchment and vellum and stuff. Then you can reuse it for, in this case, church land holdings in Florence.

And then there's these little strips that you find sometimes that you weren't even sure are music. But oh, those kind of look like quarter note stems, quarter note stems, and stuff like that. And this is what was left for me. So I became a music fragment scholar. And my things are all on fragments of polyphonic music from the 14th century.

So I wanted to look at these things. And how to move forward? One of the things back then with no computers is just spend a lot of time. So one of the things they were really good about then is that they-- who was told that in a music class, when you write a quarter note-- that's a very bad one-- when you write a quarter note, its stems should be exactly an octave long?

And so therefore, it's kind of exactly like this. I don't know if people are still taught that. Well, they were very good at that. So you can figure out, based on the heights of the stems, what notes must have been cut off there and where they are. And you just stare for hours and hours and hours of other books until eventually you find one where every one of the stems lines up. I had to squish this by about 10% to get it, but things like that.

Or for something like this, there's three parts. They wrote the parts separately back then, kind of like, I don't know, iterating through the recurse method and not the flatten method. And so part one, part two, part three. And so you just try every combination of clef and time signature and key signature until you finally find where things line up. Then use your knowledge of music theory from the 14th century to fill in the missing parts. And eventually your friends record the piece for you.

[MUSIC PLAYING]

Then when the notes get big, it's what we actually know.

[MUSIC PLAYING]

So anyhow, that was some of the things. Then the first place where computers really started coming in was about when we started getting nice color photographs, before everything was black and white photocopies and stuff.

And I, with some other people and some other people independent from me, all came up about the same time with a technique that if you measured the colors of the overwriting and the colors of the underwriting, the ink would probably be different. So you could remove one and bring the other up. And so we had these various computer techniques for getting rid of the overwriting and getting to the music underneath.

But everything after this was still going to be done by hand. So I started thinking about ways to use statistics and computers to do various things. So about this time we started creating music21. And if you ever wonder why there's some things that are obscure, music21's sole purpose was to try to identify music fragments like that. And everything else that this class has come about has just been of a side project.

So while I was building that, I switched into a statistical mode. So we'll get back to music21 in a bit. But first, I want to think about some of the things that were coming up with all these fragments.

And I was trying to figure out what they were. And that meant I would need to find another piece that was like that. But the odds that another piece would exist that matched up with the music, I was told, were really, really low.

In fact, a very famous article by somebody who's very famous in my field, Nino PIrrotta, "The Oral and Written Traditions of Music," said that the surviving music from the Middle Ages is the tip of the iceberg. And that metaphor ended up being quoted and reused in a whole bunch of other articles that you don't need to see here. So this became what everybody knew.

So if you're thinking about music from 600 years ago and you think about all the things you've lost in your 20 years of life, or however long, what percentage of things would have survived from 600 years ago? Or would be lost? Would be lost? About 90%? 99%? 99.9%? These are all numbers that I found people said. And then somebody said basically everything. So I have to convert everything to a number. That's pretty close to basically everything for me.

Let me skip over that. We don't have all the time in the world. So I started to think about some statistical methodologies that could help in this. And one of the things that we have in some cases-- this is a table of contents for a book of music from the 14th century, where the book doesn't exist but the table of contents does.

So I started looking at, OK, here's all the pieces that do not exist. And then here's all the pieces that do exist. And starting to look at the ratios. And in this particular manuscript, it looks like about 68% of them survive. So that was the first thing that maybe the 99.9999 number didn't exist.

And then I looked at there's a bunch of poetry manuscripts. So this is one from about the time of Dante or Chaucer by a not very good poet named Prodenzani. But one of the things that makes him a bad poet that's really good is that when he gets into something, when he's talking about a feast, he will list for the next 50 lines everything they possibly ate.

And when it gets to a concert, he lists with the viola, [SPEAKING ITALIAN] We played our May Day songs. We just had May Day. So these are the May Day songs, like the rose that never changes its color, [SPEAKING FRENCH] sweet taste, and so on.

And these all look like names of songs. And so we can look at, for each one of them, do we have a song by that name, or do we not? And so we have how many are known, how many are unknown. And we eventually get to a number that looks a little bit similar to the last one. About 71% survive.

So I was kind going through this. The chronology is a little bit wrong. I had to justify this later. But first I was thinking about how animal biologists, ecologists, count animals. Has anyone taken a biology class or something where you do population estimates?

So one of the ways that you can do it is you want to know how many deer are in a forest. So you go, and you capture as many deer as you can in a month and a half. You come back smelling pretty bad and stuff. But every time you find a deer, you tag its ear or some way, in some way that you can identify, this is a deer that I saw.

And you come back several months later or whatever, and then you count deer again. And the percentage of deer that were tagged in the first time give you some sense-- or the inverse of it gives you some sense of how big the population is.

If you tagged 100 deer the first time, and you come back later and you only find-- then you capture 100 more deer, and you only see two of them are tagged, you know there's a huge population of deer out there. Similarly, if you captured 100 deer and the next time you capture 100 deer, 98 of them are tagged, you probably saw most of them.

So I started thinking about, well, if a piece appears in one manuscript, we'll tag that piece. And then we'll see how often it appears in another manuscript. So I started coming up with formulas to do this type of thing. And I realized that there's an equation that tells us the number of missing pieces.

All you have to do is you take the number of pieces you originally had. And you multiply by the probability that any piece would be missing. Tell me if I'm wrong. You're the MIT students. Equals the total number of missing pieces. And it's a super great, super simple formula, except that it's a formula with three unknowns. We don't know any of these three numbers.

But we can kind of figure out some of these things. So we can look at, well, what's the probability that a given piece would be in a manuscript? Now, we'll start with the assumption that every piece is selected randomly. So we start with how many pieces are there in a manuscript? So just like a collection, like a box set of lots of tunes.

How many pieces are there in that manuscript? We'll call that r sub i. And you divide it by the total number of pieces that were originally copied. So we know how many pieces are in a manuscript because we can count that. That survives. But we don't know that, so we just move our unknown to another place.

So we can also figure out what's the probability a piece is not in a given manuscript. What's the formula? Just see if you're awake. If this is the formula it's in the manuscript, what do we do? 1 minus. Good We're all on the same page there. Great.

So what's the probability that it's not in two manuscripts? You just take the first probability and you multiply it by the second one, in the assumption that these are independent. And so try to keep ones from different places, whatever.

So then what's the probability that any given piece is not in any of the, we're at that point, 85 surviving manuscripts? Well, you just have something like this. And that gives us-- takes-- our two unknowns. It gives us the probability that a piece is missing.

So we can go back to our original equation and plug that in. And all of these have N as a denominator. So now we've at least reduced our giant-- we've reduced our master, our main equation, into two unknowns. So we're getting a little bit better.

Then we thought, how many pieces were missing? Well, the number of pieces that were missing is the number of pieces there originally were minus the number of pieces that survive. So we can count the number of pieces that survive. It takes a long time, actually, to figure out every single thing because people count differently. But we can kind of reduce this entire equation to one equation, one unknown.

And at this point, I was thinking I must have screwed up something really, really bad. I have a music degree. I don't have a stats degree. This seems like a pretty fundamental equation to have. And it turns out people had been talking about equations like this a lot in the '60s or so. But back then, it was all about, what's the way that you solve an equation like this with 85-- it's an 85th power equation, or something like that. And so all of it was on the high math that I don't understand that can get you something to this answer.

But this was my first computational thing. It's in a language called Perl, where if you know it, you're lying about your age. It was the big-- it was the language that built the internet in the '90s and things.

And one of the things you can do is you can look at-- guess various numbers and look at the error between the left side of the equation and the right side of the equation until you find the number that comes closest to balancing it. So you just kind of try everything, and you multiply everything out.

And since this isn't really a class about that, but it turned into saying that the percentage of missing pieces varied by genre, by type of piece it was. The interesting part is that the mass, the music written for sacred services, has more missing than the ones for everyday dancing and lovemaking and all that stuff.

And I think that's because over time, nobody could read this music. So the only reason somebody would keep it around is for the text. And if you're running a church, you already have lots of copies of the text of the mass. But the song about we all went off into the woods, each one as a pair, and we all came back as three, that, people still keep around. So those sort of body songs have the least missing.

So it turns out that our best estimates now are that about 25% of the pieces, maybe 50%. I don't know, maybe 60. Maybe I'm off by a little bit, maybe by a logarithmic order of magnitude or something. But we can actually really say something about medieval music based on the fact that much of it survives.

And I was very, very scared that somehow I had made a math error and stuff until there was just a math dissertation at Yale that verified all the work and stuff. So I'm finally, finally not worried I'm going to have to do a retraction of one of my biggest papers.

So now we're going to get back to these little scores. And once I had convinced myself that a good chunk of the time, 3/4 of the time or something, we're going to be able to figure out what music is on here because it's going to be in some other piece we already have.

So there's this collection that I mostly use. It's called *Polyphonic Music of the 14th Century,* and we have it over in the Lewis Music Library. It's beautiful. It's expensive. And come in these gorgeous green volumes. And there's a lot of them, all the music from the 14th century. It's about 20-- it's 25 volumes in this series. And then there's been some others since then.

So this gives you a sense of what it is. And now, MIT was able to-- I'm going to skip some of these. MIT was able to help me get some of this into a computer. So we'll get onto that in a little bit.

But what if you want to do-- if you want to find if any of these pieces has a match in another thing? Because somebody has written down a lot of these little fragments, or they've put all the pieces in, and I wanted to see not just where any of these little fragments there, but at this point, are any pieces quoting or copying from or plagiarizing from anybody else's music?

So what you have to do is you take all the songs, and you compare each one of them. And by N. Sorry, that's not enough. We've got to zoom out a little bit. And there we go. Yeah, so it's something like that. That's actually just comparing each of the volumes. We actually want to compare-- each of these volumes has about 150 pieces.

So you have to do the square of the number of pieces. Divide by 2 because it's symmetrical. And so you have about 3.1 million comparisons. And that's the reason why nobody had done this before.

I'm actually really lucky. Our best estimates, which I don't know, is that 14th century music is about 10% of the size of 15th century music. And every successive century is about tenfold more music. So I have algorithms that I've devised that I think will also work for the 15th century. And maybe if I ever had a huge research budget, I could do the 16th. But it's going to be up to you all to figure out what's the algorithm that can compare every piece of known music to every other piece and get that. And that's why I care a lot about similarity.

Sorry. I got ahead of myself. So the great scholar, Anna Kathryn Grau, created the Electronic Medieval Music Score Archive Project, or EMMSAP. Slide's a little bit out of date, but back then, it had about 1.3 million notes. Now we're close to 1.8 million. And, of course, you know this one, right?

So what you do is you take all the music, and you take music21. And then it does something like this. And that's how it looks when you run the matches. Now, you all know that this is what you do to get funding for your project. And that's not actually how it works, right? Only in CSI.

No, the way it actually works is you go, and you take your scans that you've done. And how do you get them scanned? You spend all day putting them on a scanner. And then enter every one of these pieces into mostly MuseScore, Sibelius, Finale, whatever it is. And then we exported everything as musicXML. Some people like-- there's a specific output for musicology called MEI, which is pretty similar, but I use musicXML.

And so then we have here's the musicXML for the first three-- first four notes of this one piece called-- sorry, I'll just go back called "Onques Ne Fu," which is completely only copy of it in the world.

Then I wrote some programs. Python, music21, things just to take every 31-note little segment-- I don't know why I ended up with 31-- and just automatically chop it up so that there's these little segments, things.

And ideally, you would chop it up and you would slide your window over one note at a time. But there's a trade-off between getting the best similarity and also not having your database explode with internal quotations. So we did this and had it go through everything here.

And then once we had chopped up the pieces into these little 31-note segments, then we did what Keith was talking about, a viewpoint or a hash or an encoding system to try to reduce the complexity of it. So the complexity of the system that I used, you're going to laugh at.

If a note was the same duration as the previous, it got the ASCII letter A, B, C, D, E, F, G. I dropped sharps and flats because different copies of the same piece in the Middle Ages sometimes write it as B-flat. Sometimes they don't write a flat. You're just supposed to hear that. Well, obviously there was an augmented fourth. I'm going to sing that flat so they don't always put that in. So we get rid of accidentals.

And then if the note's longer than the previous note, then we use H through N. And if it's shorter, then it's O through U. So P is a B that is shorter than the previous note. You just go through and put everything in like that. And the first note, I always just assume it's the same.

So we take all of these things and put it into a giant database. It's not that big. It's about 188,000 of these little segments at that time.

And so this is the segment for the opening of "Onques Ne Fu." I always think it's kind of unceremoniously stored there, like the Ark at the end of the *Indiana Jones* film, if you've ever seen it. You just can't really tell where it is.

So then we have to compare every one of these 188,000 sequences against every other one. So we take the first one, the "Onques Ne Fu" melody, and just some random other one. And we can figure out what are all the things we have to do in Earth Mover, Levinstein, or some other distance thing in order to transform the first one to the second one. So you have to-- well, you insert a whole bunch of things, and then some things are the same. Delete, whatever.

In the end, you get a ratio from 0 to 1 of how similar they are. I always like to store integers in databases. It saves a lot of space. So I just multiplied by 10,000 so I get a number between 0 and 10,000 to store.

These ones are 17% similar. These are actually a surprisingly bad amount of similarity. Random pieces tend to be more similar. So this is not a match, right? Nobody's going to say that those two lines are the same piece. Slightly different. No.

But that's OK because we do it 188,000 by 188,000 times, or 35 billion. Divide by 2, since it's symmetrical. 16 billion. 18 billion, right? But then I went through and stored every single time there was something that was more than 50% similar. So there's two IDs and the ratio of how similar they are.

And over 50% similar ended up with 814,000 things. I'm never going to look at those. That's just way too many. So I started going to-- at around 65% similar, I started to suspect these tended to be pieces by the same composer or quite often, if we knew or I suspected they would be, or something like that.

At 70% similar, you're like, wow. These aren't the same piece, but they're a lot like it. But definitely over 80% similar, they're always the same piece. And how can something be the same piece and be 20% off? It could just be that they were sliced at different points. So they're the exact same thing, but the first third of one segment is in some other one.

So this is what the actual workflow looks like. I used to use Eclipse. And just running through all these pieces. And occasionally it finds a match, and it puts it up. And I get to see, OK, does that look like that? Yeah, that one kind of looks like that one. And sometimes we already knew them. The computer does not read what pieces we already knew were quoted.

But we ended up with some interesting things. This is a piece by a guy named Johannes Ciconia. This is unpublished yet. Don't put it all over everywhere. A guy by the name of Johannes Ciconia, one of the most important composers around 1400. And one of his pieces, it's a part of the mass, patrem omnipotentem. And it turns out that a big chunk of it was taken from an unknown-- an anonymous, never recorded by anybody, German song.

[SINGING IN LATIN]

Let me see if I can do the-- Here we are, if you've never read it.

[SINGING IN LATIN]

So we can find that there was a lot more copying than we ever expected. This is one of my favorites. Here's a French piece, an Italian piece. Nobody ever suspected they had anything in common because whoever-- and I don't know which direction the copying happened-- whoever did the copying did a good chance of covering their tracks. The colored parts represent the same music. And then down here, stole something from up there. And then up here, we stole something from over there, and so on. So it's just two that are pastiches of each other.

And then if this were a musicology class, people would actually care about which one this-- what all these other things. But eventually, we got to start reading these erased pieces and figuring out where they came from.

So that's the big reason why I created music21 and why you will find that there are lots of weird little things that are only useful for music, for medieval music, hidden around deep inside the various libraries and stuff. So that's what I do when I'm not enjoying being in your company. Any questions? Or anything you want to see again?

There are, by the way-- lots of places do want historians who can do this type of work. So if you're all interested, humanities in general. But you're like, but I also want to program and stuff. People are dying for things like that. Yeah?

AUDIENCE: So what made you want to expand music21?

MICHAEL CUTHBERT: You all, your predecessors, people from about 10, 15 years before you from here at MIT. It just felt like, people wanted to do interesting things with composition, with analysis. Everybody wanted to solve their first-year theory homework by computer. And why not? If you can program the computer to do it and it's not a complete black box AI, you're probably teaching yourself how to do it better also.

So it was things like that. And over time, I've taken out a lot of the most egregious that only a person working on 14th century Italy would want to do this and put this in a separate project. So there's no longer an automatic "change medieval notation into modern notation" tool. You have to install that separately.

But yeah, it's just felt like I was having too much fun. And other people wanted to share in it, and that meant eventually docs would have to happen. And then a very generous funding organization called The Seaver Institute selected music21 to be MIT's representative project for their funding organization for three years. And that allowed me to do a little bit more than one person can do at their desk.

And now the thing that helps me do more than one person can do at the desk is community of a lot of people who keep adding to it and doing cool things with it. So thanks. Other Qs?

AUDIENCE: Why is it called music21?

MICHAEL CUTHBERT: Course 21M. I wanted something that connected itself to MIT in some way but wasn't MIT music because your trademark office would probably get against me. And also, I was a second-year professor without tenure, so I might have been denied tenure and needed to be able to take it with me if I went someplace else.

And then somebody later was like oh, it's the thing for the 21st century? I said, I actually never thought of that connotation. If I had, there's no way I would have called it 21 because that sounds very A, arrogant, and B, a little obsolete in 2006, when you're already in the 21st century. It's perfect for 1980 or something. Yeah.

AUDIENCE: What was it called before music21?

MICHAEL CUTHBERT: It was called PMusic first for Perl music when I was writing it in Perl. And then the fact-- then as soon as the Intro to CS class switched from your older language scheme to Python, I was like, well, if all of MIT is going to move to a new language at the same time, then it'll be easier for me to learn a new language than to have the rest of my life teach my students how to program in Perl. And I'm very glad I did.

The other thing is the work on it began in 2004. The work really began around 2006. And at the time, I predicted that it would have a 20-year lifespan, that it would be five or six years of getting people to notice it and start using it, seven or eight years of it being the main thing out there, and five to seven years of the decline while something else better came out.

And I'm still looking for that something better. You guys can do that. There are things we've learned from using it that I think have made me realize that there are some pretty poor design decisions, such as having to copy everything all the time and things like that.

And so I'd really love it if somebody else would create something new and better. So whoever's here or watching it, you have my blessing to try to knock this off its shelf. And maybe the new thing will have a lot of machine learning built in or something to be able to do better analysis. Yeah.

But Python has been enduring a lot longer than I thought it would be. It's been 15 years as a major language, and it seems to be still growing, so that probably will stop.

AUDIENCE: Julia is next.

MICHAEL CUTHBERT: Julia is next. People have been saying that about Julia and R. There is a Julia port of music21, if you'd like that. I don't know how good it works. So that's a perfect transition.

We can keep going more to a second topic. So this one, get out your laptops. We're going to have-- because I couldn't have scripted a better transition to-- OK, so now more grab bag. And these are more requested from previous years than anything. Sorry. I can't stand all that dust. Oops. Did pushing all this shut it down? No. Here we go. OK.

So here are some of the things. Go ahead and go to this website, m21sand2. If you leave off the 2, you get to use a three-year-old version of this. You should be seeing something that looks a bit like this when you get there. 21m-- 21-- m21sand2. m21sand2. sand Sandbox, the place where you play around.

So the thing I've always hated most about music21 is that you can't see anything. There's two things. You can't see anything. You can't hear anything unless-- we didn't even have Jupyter notebooks or something back then, so we didn't have that. And number three, it's so hard to install. That's getting easier, but still.

So I thought that something where people can start playing with it immediately online would work. And so I've been working in parallel for the past five or six years, there's a TypeScript or JavaScript port. Who has done JavaScript here in some ways? OK, a little bit more than half.

So the big difference if you're new to JavaScript is that you're going to put new in front of any time you want to instantiate something. And I've imported music21 without importing all of the things. So you're going to have to type that quite a bit.

But what's kind of nice is you can go through and change a note and immediately see the stream there, or octave, things like that, or put everything down. What happens if we put everything down here? Doesn't it automatically give me the good clef? No. Did I put it in? OK, well, we can put in a proper clef. s. So we'll append to the stream a new music21.clef.BassClef.

And now we can see our notes again. And by default, if you click the thing by default--

[NOTES PLAYING]

There we go. OK, well, I've lost all my changes, but you can see. And so some of the things you can do is create interactive little systems where you're scrolling things up and down, where you're hearing everything. But it's still absolutely a work in progress.

So if you do analyze key, you get analyze is not a function. So there's still a lot of things missing there. But there's something really nice to being able to manipulate with scores in a browser.

The alternatives that people are doing right now tend to be about actually using HTML elements to represent notes, now that you have custom elements. And that lets you use your browser's built-in getElementsByClass and all these things to move around a score. And in theory, it should be substantially faster than this, or at least more powerful.

I don't know if anybody wants to-- you can play around with any of this for a little bit. You can make this a dotted quarter. What I hate is duration.dots equals 1. And nothing's going to happen until I get rid of this analyze thing here. Yep. Play it around. Give it a 7/8 time signature.

And then-- here let's do this. Let's create p equals stream.Part. Oops. I'm totally-- just being in this class with you all, I can't not be in Python mode. So I'm just going to keep typing Python and then correcting it to JavaScript until it runs. But we can do something like for let i equals 0, i is less than 4, i plus plus.

And then-- oh, this pains me as a Python-- as a Python person to not have indented code. But actually do we have to call that? We have to call that Measure. Music21j is a lot more strict about such things.

Did I-- Hmm. What did I do wrong? It's so hard to tell without indentations. But this editor doesn't allow me to hit Tab to indent things.

Always make your life easier. Get rid of that so you can see everything on one page. Anybody know what I'm mistyping here? I have a fully baked version over here somewhere on another tab.

Here, we'll go over to the fully baked one that already works so you can play around with it. Just saying first note of each measure will be a random number. And as we go through, we'll just change them. And if I make any changes here, it will automatically give me a new random number.

I thought I had it. Oh, this. OK, I think I have to copy and paste. What I think happens is-- there we go.

[MUSIC PLAYING]

You can do overlapping things.

[MUSIC PLAYING]

Anyhow, so it can be kind of fun for a live coding environment if you don't care about the quality of the music coming. If you don't care about the quality of the sound because it's not really optimized for sound quality, but you could put it out through other things.

Let's grab bag topic number one, how to do interactive-- how to do interactive programming. So if you can do everything-- sorry, John. You just came back in, and I'm going to just say, you're the only one. Has anyone else taken Interactive Music Systems? Oh, you are?

OK. So you have-- yeah, so you can put everything into-- you guys usually use Kivy, right, as your toolkit? So you know there are things you can do on the Python side that connect up music21 to things to get better sounding interactive things. But I often find sometimes just the easiest little thing can be all you need.

OK. I'm going to close these things because otherwise it will eventually just start playing again. And we've just enough time for, I think, for one more topic, and then I might finally keep my promise and let you all go a little bit earlier.

OK, so the last thing I want to talk about a little bit is music visualization. And we're going to listen to a piece of music by a contemporary Estonian composer, Arvo Part. And I'm going to ask you afterwards to say something about the structure or the form of this piece and how it works musically.

First, it's a piece for choir, and so it has a text. It's actually high, high medieval today. It's based on a Gregorian chant. It's sung in German translation.

[SPEAKING GERMAN]

And the German texts can be slightly important for that. So it's about a two-minute piece and wrong time of the year for doing it. But with that, here we go.

[MUSIC PLAYING]

1988, I think, was one of my favorites. Anybody catch anything about the structure? You have seen the entire score, every single note. So we can get something. Yeah?

AUDIENCE: The measures always switch between 1/2 and 3/2.

MICHAEL CUTHBERT: The measures always switch between 1/2 and 3/2, 1/2 and 3/2. How many sections did the piece have? Yeah?

AUDIENCE: Into three.

MICHAEL CUTHBERT: Sure. Nice. That's actually correct. How many people knew that? And how many people were like me when they heard this once, twice, three times, and didn't fully catch the structure? I was-- yeah, yeah. Fantastic, fantastic.

Usually-- so yeah, there's some things. And what voice is most-- what two voices were most melodically active? Yeah?

AUDIENCE: Middle two, I think?

MICHAEL CUTHBERT: Good ear. Good ear. Y'all are really getting-- we should have been doing more music analysis in class because you're getting a lot of things I thought were going to be extremely hard. And at least throughout the room, we've had correct answers.

But I want to spend a little bit of time on music visualization. By the way, there are a few-- out of almost exactly, there is one measure of 5/2. But yeah, there's a couple other things, but yep.

So are we up there? Yeah. Go ahead and download. There is the XML file, musicXML file, of this piece on Canvas. And take a minute. It's under copyright, but I'm going to assert a claim of fair use for us to be able to analyze with it. Should show you just the first page if you're in Jupyter, and more than that if you're not.

So some of these things might be helpful for what we're doing. So a little bit more. We're going to be going through the various ways of visualizing.

I don't think we've maybe I've done a tiny bit of visualization, but mostly just on slides. I don't think I've shown-- are we feeling in a good place to move on now? OK. So we can do all the things that we normally would do. We can get the key of the piece, which I believe changes quite a bit.

I think if we do measures 1 through 20, analyze key, we get something very different. No, we don't. Still get-- interesting. Just the key signature changes. OK. But now we'll get into some visualization things.

So the easiest thing you can do with a stream is to do something called plot. And plot just gives you whatever you-- a standard, ready-made-- oops. That's why I never do this here. Ready-made image.

Sorry. It doesn't zoom the image in here, but you might be able to see something like that. And so you can see a relationship of the various parts colored in different colors and which ones are singing what.

So here are the three-part-- the three-fold relationships. Sorry, I'm going to have to zoom this out a tiny bit so we can see this again. The three-part relation-- three-section relationship of the piece becomes very clear, right? You can see that something happens at the beginning, this increase up here up until a certain point. And then a lot of things change.

So what measure does the third section begin on?

AUDIENCE: [INAUDIBLE]

MICHAEL CUTHBERT: Yeah, somewhere around there. So you can-- 45, I think, 46, 44, 45. So you can see something substantial changes. And here's what Jason pointed out, that the inner voices in the second and third sections are the ones that are very, very active.

So where does the second section begin? Measure 13 to measure 45. So let's get this. sec2 equals arvo.measures. Let's say 14 to 45. 44, we'll say sec2.plot. I think the default is piano roll because that's what most people want.

So we can look at Section 2. And we'll say section 3 was measures, what, 45 to 76. Section 3 equals arvo.measures 45 to 76. We can plot that.

And there's a lot of similarities here between section 2 and section 3. What's the biggest difference musically that you can tell from these charts between section 2 and section three? You'll probably want to go all the way back up to your bigger roll.

Over here? This side's been mostly quiet. Yeah?

AUDIENCE: Section 2. Section 2 and section 3 close, but section 2 is up.

MICHAEL CUTHBERT: Section 2 and Section 3 are almost identical, but section 2 is up. Good. Can we figure out about how much up it is? Look at the key on this side. Yeah, the E to the A and the E to the A So it's about a fifth off in a lot of things, except for on the extremes.

But what about the notes in between these moving parts? Where are they off? What's the relationship there?

AUDIENCE: An octave up?

MICHAEL CUTHBERT: Yeah, about-- exactly an octave up. So we have things moving. Now, we'll go-- I'll open up the score one more time. I know why I wanted the score open, so I could keep going to back to this. Let me get a-- yep. There we go.

So I'll just put this in front for a second. sorry, now I've lost my train of thought. How easy is it to see which parts are static and which parts are in motion in this? Quite a bit harder.

So what kind of motion is happening? We've identified that it's the inner voices that are in motion. What's the pattern? Now go back to your visualization and see if you can figure it out. In any of the inner voices.

Maybe it's easier if you just look at one of the sections. So you can zoom in a little bit more. John, do you want to?

AUDIENCE: They move in parallel sections.

MICHAEL CUTHBERT: They move in parallel first, yeah. That looks good. And within any given voice, what's it doing? Is it playing, "Mary had a little lamb, bop, bop, bop, bop, bop, bop? Is it jumping around a huge octave and then coming back down? Adam?

AUDIENCE: They're just playing the same note.

MICHAEL CUTHBERT: They're just playing the same note. Not all of them are playing the same note. These ones are playing the same note. But what are the other voices doing, these inner voices? That go all the way up and then all the way? Yeah?

AUDIENCE: Are you saying they're repeating?

MICHAEL CUTHBERT: They're repeating, but what do they do? What's the pattern? Yeah?

AUDIENCE: Stepwise?

MICHAEL CUTHBERT: Stepwise. Good. It's all stepwise, which is the hard part about the piano roll, right? You would want a diatonic piano roll at some point. I think we have that somewhere. And it's all stepwise. And it's even more than that. Yeah? Go ahead, Vincent.

AUDIENCE: Neighbor tone's around a certain pitch, just like the top voices around the C.

MICHAEL CUTHBERT: They're all neighbor tones around a certain pitch, yeah. So you can-- top one, I can't reach, so we'll do the middle one, I guess around E. And is there a pattern that you do three ups and then a down, and then what is it? Up, then down. Up, then down. Up and down.

So you could almost have a formula with this. In fact, one of the other Python-based music packages, which is more about creating algorithmically generatable music, is called arvo, based on Arvo Part's ability to do this. So you can download and install arvo, and it makes it much easier than music21 does to generate. Here's a motif. Now spin it out over a certain amount of time while doing something else. So we have this thing going up and down and up and down and things like that.

By the way, some of the other interesting things about the piece that we're not going to visualize in music21 are, when does it change notes on any multi-syllable, accented German word? Which I think is interesting here.

So the accent always gets the whole note. And it's always the opposite of whatever it was before. And then you can see, well, in section 2, all of the voices kind of play at once, and they all rest at once. But in section 3, the top and the bottom ones only come in at certain times.

And the visualization of this helped me in figuring out, well, when are those times? And it turns out, also very strictly generated, that it is just every other word. du Hoffnung, they do. That's two words. und Heiland, we don't sing. Next. Two words, der völker, we sing. o komm, we don't.

And there's one exception to it. No, sorry, there isn't an exception on this one. Someplace else, there's an exception on things. But I think probably set up so that just the number of words makes it so that God stands alone at the end without accompaniment. So interesting piece.

Now let's do a tiny bit more analysis on it. There's a plot that-- I could call this a cook plot. Cook, because I saw this now in the reading for Nicholas Cook article on computational and comparative music-- computational and comparative musicology.

But I first saw this in a book by somebody named Emily Dolan. So I call it a Dolan graph, even though-- and it shows you who is singing where, at what point, which parts are performing. And do you remember this from the Cook? It's slightly different graph that he had, a slightly different shape, size, but same type of thing.

What's being depicted here? Take a guess. What's that?

AUDIENCE: Dynamics?

MICHAEL CUTHBERT: Dynamics. Yeah, so it's a who's playing when, and what are the dynamics? And if you zoom in really closely, you can see that soprano is always singing a tiny bit before the other parts. But yeah, so the piece just sort of builds up to about 2/3 point, then drops off and just very quiet, everything after that.

So there's quite often, I think, for people who can read music notation, graphs and plotting, visualization isn't too much of a help. But for getting a sense of the form of a large piece very quickly, these tools can be very, very helpful.

We can see, what notes does he use the most? Well, obviously those that are going to be-- that those sustained notes are going to come quite a bit. But we can see that of the various pitch classes, basically, there's a reason why it was coming out as A major, as our analysis. Because A, definitely winning. E, C-sharp.

I guess every-- surprisingly for a piece that sounds so consonant in many ways, every pitch class is represented at least once there. Let me see which-- I have a whole bunch of ones that are maybe a little bit less interesting.

But we can look at certain pitches use different notes more. So we can plot a scatter plot. And then what we haven't seen yet is the second and third element here is quarter length versus pitch.

I think there's other things you can do. You can play around with octave and stuff. So a plot gives you a ready-made take in a score and get some kind of analysis. There's two slightly lower level things if you want to-- if you're thinking of doing visualization in there. One, the graph, which is underlying the plot, which is what you're trying to do and what data you want to extract. So it's a feature extraction again. And what do you want to put on each axis? So there's axes that are defined that you can--

I don't know, we can do-- let's plot pitch against dynamic. You know What? Yeah, dynamics. Does that work?

You know what? Who puts pitch on the x-axis? We'll put pitch on the y-axis versus dynamics. Swap that around. That's why I was having so trouble.

So you can see that the-- oops-- that the lowest notes in the pieces only appear on the quietest things. And you can also look at this terraced initial motion of the piece as the notes were going up, both going up in pitch and getting higher in volume.

And then the very highest volumes are scattered all throughout, but everything down here shifted up a little bit, just as we saw.

I think the best thing to do would be now that we've thought about the piece, thought about its structure, listen to it one more time. Maybe with the-- we'll listen to it with the piano roll and see if we can-- oops, not that one. The big one. And see if we can hear the various parts of the piece just as well as we could on with the score.

Words are harder to tell. There's no plot for that.

You're mostly colorblind. Professor just realized that I used a version of the plot that does colors not based on the part but just as a way of being able to see which pitches are which. So I apologize. I made a mistake and said you can tell the lines by color. But I have to strain to see color, so sorry about that.

But I did, just while listening to it, make one observation. This interesting line that the E is the very first thing that you hear. And it just after a little absence comes back. And it's one of the few lines that's so prominent throughout there. So maybe E is the true root point of this piece.

Well, thanks, everybody, for going through grab bag. I have more topics that are interest to me, but anything that's of interest to you, let me know.