[RUSTLING]

[CLICKING]

[SQUEAKING]

[CLICKING]

**MICHAEL SCOTT ASATO CUTHBERT:** Hello, everybody. We're going to be talking today about something that sounds really esoteric, really confusing, vocabulary reduction. How do we go to vocabulary reduction? How do we reduce our vocabulary? Are we talking about see Spot run? What does this mean?

It really is about something that we do every time we do analysis whether we do music analysis or textual analysis or we're trying to think about what really happened on that last phone conversation I had with my partner, friend, or something like that. We're trying to reduce the number of complexities down to something smaller and more manageable.

And one of the places where we do vocabulary reduction quite a bit is in Roman numeral analysis. We don't label everything a Roman numeral. We tend to label the things depending on who your teacher was. Maybe she or he said we only label things that are really chords.

What does that mean? When we're thinking about how we do this as humans, we don't really have to think too much about how we do this. We just have to get better at doing this. But if we're going to train a computer system to be able to tell us only the things that are really there to reduce the vocabulary of a score into something more manageable, then we're going to have to understand this process quite a bit better. So that's what we're going to do today.

I'm going to move over to the Jupyter Notebook, and today I didn't have time to make every single thing a little bit really polished. So you'll get a sense of how I program, how I make up assignments for us, and how some of the wrong turns that happened along the way. Maybe it'll be useful. So let's go over to Jupyter.

We're going to start by importing from music21 the corpus method, and let's load up a piece by Clara Schumann, her *Piano Trio,* slow movement, one of my favorite pieces. And we've looked at this once in class before, and it's going to take a little bit of time. It's a little bit of a slower thing. You might need to wait here.

It's a piano trio, but it begins with just the piano parts, really lovely section. Let me find a YouTube video and place that in here.

[MUSIC PLAYING]

And you can click on that or just Google it, find it in there. Anyhow, the next eight measures we looked at once in class, and we can see it's when the violin and the cello enter and the piano switches to something else. The violins play the same melody, and we looked at it because of those grace notes, which if we can highlight that there. There we go, and they-- yeah, they were a problem for some of our representations of encodings.

So let's just extract the piano part from here, and you can see some of the mistakes that I've made with my own music21 package that this happens. I'm like, oh, I want to extract two parts.

Well, first off that should be two to four anyhow. But in any case, it's given me a list of things. So I've made two mistakes. That's not giving me a stream. It's giving me a list. I'm going to--

Now I'm going to do this again. We're going to import the stream. We're going to-- to create a new score, not part, for putting things in. And we're going to put in-- gosh, I cannot type piano after all these years-- we're going to put in part two, which is the piano right hand, and we're going to also put in part one, which is the piano left hand.

And in order to make this run faster, we'll just take the first 16 measures. I put measure 0 just out of habit in case there's a pickup measure. There isn't in this case, so I could have really started with one.

So now we're going to look at just that little excerpt, just the piano part, all that to get rid of some empty rests. But that's OK here.

So we're going to play with just creating the chords of this, and we'll remember the chordify function, which is a fun part of music21, and we'll show it. And because we're taking a treble clef and a bass clef put together an instrument, this is always a little bit wacky. You get used to counting obscene numbers of ledger lines when you are playing with chordify And there were times in my life when I could actually eyeball that first note and say, yep, that's G. I got it. But those times have passed.

So instead we're going to recurse, go through every single element in every measure and everything, and get all the chords from that. C chords is now a part. And we're going to take every one of those chords which we call C and put it into closed position and say that it's going to be an octave four. And in place equals true, we're going to replace the old chord, and now we're going to show that the-- are the same thing. And now things look quite a bit better.

So, yep, there's a-- think we compare back up if we look at same page. Yeah, a little bit. Yep, that's a D down there. That's a G. So it worked. Worked pretty well. We'll zoom back in again so that we can see-- great.

Now what we're going to do is we want to label everything that's happening in this. So first we're going to analyze the key. What key are we In. We're in G major. Yep. You'll see that this is called *The Piano Trio in G Minor.* But this movement is in G major. Great.

So we're going to iterate over every one of these chords again, and we're going to create a Roman numeral for each of the chords. Actually we're going to import the Roman package from-- Roman module from music21, and there's this cool thing that you wish you had in your last time you were programming your last problem set, Roman numeral from chord. So we're going to take the chord and say that we want the Roman numeral in G.

We're actually going to create an object called G underscore major, which is a variable which is the key of G object, and the reason for this is otherwise we would have to create a new key object for every chord and over time that becomes slow. So what we'll say is for our original chords, we'll give them the lyrics which are equal to the figure of the Roman numeral. So whatever the Roman numeral is, we'll assign that to the lyric. You can see it's not the fastest thing in the world, but now you can have it and we can look in.

We can look at those odd chords in a little bit. You can go ahead and pause the video if you want to look at them in more detail. Have a nice laugh. But I'm just going to just get a quick excerpt, the first four measures, that first line there, just so that we can compare without having to move our eyes up and down a lot.

So now we'll compare this to the original piano part with the left and the right hand and just see do these chords look about right. Yep, you see, there's that weird thing that that second chord isn't really 6/4 chord. It's really that that low G is self-sustaining, and then it's at a four chord or is that 5/4/2 chord? It's going to be a little bit difficult because we want to make sure is that an appoggiatura or is that a chord tone. It's a little bit difficult.

In fact, this would be the challenge problem to be able to take something where we have arpeggios, melodies, and things and do vocabulary reduction on a piano part like this. This is still one of the big unsolved problems in the field, and being able to make progress on how do we reduce the number of chords of Roman numerals and analytical things to only the, quote unquote, real chords on a complex score, on an orchestral or a piano score. Still, one thing that you could get A plus plus. You could get complete published paper out of being able to do that.

For now, we're going to-- and there are some great ways to do that. We'll be looking at a dissertation by a guy by the name of Christopher William White at some point for people who are very interested in such things.

We're going to be applying the concepts of vocabulary reduction to box four-part chorales because these this is music that tends to have all of the harmonies changing very often, and not having a lot of the arpeggiation and things that we were just seeing. You're going to be working in groups and I'll talk about the problem set later, but I want to say that there's two different approaches that you could be using and perhaps more to do this kind of assignment. What I'll be doing in the rest of the video is an approach where you start by labeling every vertical element, and then you start removing ones that don't seem to actually be chords. This is a great way of working.

Another way of working that resembles perhaps a little bit more closely how we as humans work is to start by identifying non-harmonic tones and putting them all out in various places and then removing the vertical chords that are not real chords perhaps by also using the various Roman numeral sections.

To give you a sense, I want to show a little bit one particular Bach chorale. We'll look at the opening here of a four voice piece, and you can see moments that, oh, that looks like a passing tone. That looks like it's going to be a little bit-- yeah, just on its way someplace. Here it looks like a neighbor tone.

This might be a suspension to here, or this might be a truck suspension and so on. So we can do it by hand, but then how do we train the computer to do it? So that would be one type of question. And then once you've trained a computer to identify these tones, you can create a version where you remove them and chordify.

After this comes an essential step of identifying where various key changes might be. So regardless of whether you start with looking at finding passing tones, neighbor tones, and so on or start by looking at every particular chord and removing, the end result ends up being the same, and your work can be judged by how well each various part had been identified through your process.

I would suggest if you're working and dividing up the tasks among people in your group that you start with the removal of various individual non-harmonic tones. If on the other hand your group chooses to work where everybody is working together at each stage, then perhaps the technique I explain in the rest of the video where you label chords first will yield better results.

But we can go and do something a little bit easier if we stick to chorale style compositions. So we'll go back to our favorite Bach piece, BWV 66.6, which is going to have its own problems in a little bit, but we're going to chordify it. I'm just going to cut and pasting up here and we're going to again, once we chordify everything, put it into closed position, put it into octave four, and, again, replace all the chords in place.

If you haven't seen how this technique works, look at music21 user's guide, chapter nine and chapter 10. We do that a lot there, in fact, with this piece. So you'll see some new techniques and be able to do it a little bit more slowly.

So this piece has the interesting characteristic that it ends in a different key than where it began. So if we look at the key of measure 0 to 2 using Carol Krumhansl's probe tone key analysis method, which is built into music21, and we see that it's in A major. But then if we look at the last seven measures 4-10, we see it's an F sharp minor.

So we're going to work on each of the two parts separately. So let's first look at measures 0 to 2 and get this excerpt as B chords underscore excerpt-- should probably call the excerpt one, whatever-- and we can look at those first two measures. And let's use the same technique that we used with Clara Schumann to go through each chord and get a Roman numeral from it and then put the lyrics in so we can see it. We just have to change G major to A major.

Again, you can just put the string A there, but I prefer to do this. So we'll do A major and A major and change C chords to B chords, excerpt, and of course A maj is not the same as A major. So we'll go ahead and fix that. Put that in. And now we'll take this excerpt and show it.

And we can see-- this is quite a bit simpler. We have the 1, 5, 6, 6. Yeah, things like that. This is actually a really good time to remember think about references because we haven't really talked about this. So this is just the excerpt of just the two measures.

When we go back and look at all of the chords, one of the things you can see is that the figures do sustain back to that part. In other words, the chords in the excerpt are the exact same objects as the chords in the chord part. So what you change on one of them will change on the other. But if we go all the way back to the original Bach part and we show it, we'll see that, hey, there's no lyrics anywhere, anything because there were no chords there and we just can't get back to that.

Anyhow, that was a little bit of a digression while we look at the Roman numerals here. These look like the kinds of Roman numerals I was taught. So there's not too much, too much to do there.

So let's go back to the second block excerpt, the one that's in F sharp minor and measures 4-10. So we'll go ahead and get measures 4-10, and we'll call that block excerpt B chord excerpt 2 and measures 4-10. When we show this part, you're going to see that there's one problem with all of our showing things, and that is to say that the measure numbers are not going to be kept but whatever.

So we're going to go ahead and create a key of F sharp minor and pause this for one second in the video to see that when we make a key-- a key object really important to make a lowercase f if you want to show that something's minor, and I call it fis by the way because you can't obviously put f, number sign, pound sign to mean a sharp sign as a variable because it would just be a comment. So we sometimes use-- borrow things from other languages, and so f-i-s, fis, is the German for f sharp.

So getting back here. Getting back here we're just doing the same thing that we did before, B chords x sub 2 We're getting each of the chords, change them to Roman numerals, taking that Roman numerals figure, and make it into lyrics, and now we're going to show that section. And these measure numbers are the thing I was talking about that I guess we have to add three to each one of them to get what it is. Eventually we'll be able to contribute back to open source projects like MuseScore and be able to fix this limitation. I haven't had time to do that yet. Such a great project.

Anyhow so we can go look through and see that this time there's a lot more interesting chords and some things that definitely we want to reduce the vocabulary of because I don't think all of those chords actually exist. Let's just grab that last measure of the second system here, which is measure 8, and look into that over here. Yep, so there's an A sharp and then a B and a C sharp in the bass. Yep, that's definitely the right measure.

And you can see that that chord, quote unquote, on the second eighth note is almost certainly not a chord. That is a leading-- a passing tone, and so that flat augmented chord on 6/6 or something, whatever that was, that's something we want to reduce in the vocabulary. So how can we do this?

I'm going to give you one quick way that we can do this. Let's just excerpt measure eight chord-- 8 and just get that one measure from the-- this is from the chordified version-- measure 8. And we can just show it, make sure we got the right one.

Always do this by the way. The number of times that somebody has worked hours and hours and can't figure out the bug, because it turns out that wasn't the right bug-- that wasn't the right measure. So, yeah, I'm going to pause the video for one second and go through those chords in some detail.

So sorry, I've cut off the soprano, but it's just a whole bunch of F sharps. So we can see that the first chord, 1/6, that's really from the bass. That's the chord that's happening. It's either the chord on beat 1 or the chord on beat 2.

The third beat, the bass line looks like it's doing some kind of a passing tone, and then it's moving up to that C sharp at the end. So something is happening and-- but, well, it's really unclear if anything except the first beat is something that we'd want to keep as an actual chord, especially looking at the alto line looks like it has both of those Ds might be passing tones and the tenor also looks like it's passing tone. So a lot of these chordal motions are things that are introducing interesting elements, but are they introducing new chords?

So what we're going to do in the next section is let's just get a way of saying which chords do we think have enough stability for us to label them. So we're going to look through each chord in measure 8. In this case I don't need to do recurse because measure 8 only has chords and time signatures maybe or something like that, and I'm good at print whether or not the chord is consonant-- that's a routine, a method that's on every chord object-- and then print the chord object itself.

We can see that only that first 1/6 chord is actually a consonant chord. Everything else is dissonant according to the rules of tonal harmony. So what we could do is we could create a new stream and start putting all of the chords in, and then let's replace that back into our original, or we can start copying things and putting them in. But after thinking aloud at the keyboard for a little bit, I decided I was going to just go through and let's just remove all the chords that are not consonant. So every stream has a nice little remove functionality from it.

There are some problems by the way just like all Python and many other languages lists and arrays and stuff like that that if you remove things while you're iterating over, there could be problems. But here it's OK. So we're going to remove every chord that's not consonant, and then we're going to get that first chord, the 1/6 chord, and we'll make the duration of that be a whole note.

So now when we look at measure 8, we see only that consonant chord. Ta da. And that's one of the forms of vocabulary reduction that's going to be pretty common.

So when we go back and look at the excerpt here for the first part, we don't have nearly as much to remove. So maybe we want to do the same type of thing. Well, let's just look at that first excerpt and see what things are consonant, what things are dissonant. And here we do need to recurse.

So we're going to look at this first excerpt and see, wow, almost everything here is constant. There's that one false. Which one does it refer to? What's it refer to? Oh, let's just replace it and print the figure instead-- not the figure-- the lyrics instead and see that that one thing that is not consonant is a 5/7 chord, something that we might want to keep, whereas the something else that is constant is a 3/6 chord, which we might want to remove and say, well, that's not a real chord. In this case, it is because we're about to be modulating into another key.

So we're going to be thinking about how can we do these things, what are the various criteria that we can find, and how do we do this?