Please

Ask questions through the app

Rate Session

Thank you!

@GOTOber

gotober.com

Let’s talk about Bach

BachBot

BachBot.com

0:04 0:14

1	Sample-1	▶ 2,244
2	Happy Birthday	▶ 2,091
3	Twinkle Twinkle Little Star	▶ 6,524
4	Sample-2	▶ 1,304
5	Sample-3	▶ 800

View 8 tracks

Like Share More

https://soundcloud.com/bachbot

Let’s talk about ~~Bach~~ BachBot

BachBot

Composing Bach Chorales Using Deep Learning

Feynman Liang

GOTO Berlin, 24 October 2019

About Me

- PhD Statistics @ UC Berkeley
- Previously:
  - Director of Engineering @ Gigster
  - MPhil Machine Learning @ Cambridge
    - Joint work with Microsoft Research Cambridge

BachBot: Automatic composition in the style of Bach chorales

Developing, analyzing, and evaluating a deep LSTM model for musical style

Feynman Liang

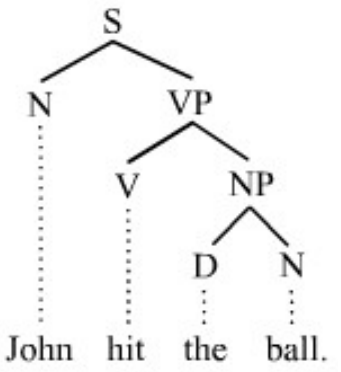
Department of Engineering
University of Cambridge

M.Phil in Machine Learning, Speech, and Language Technology

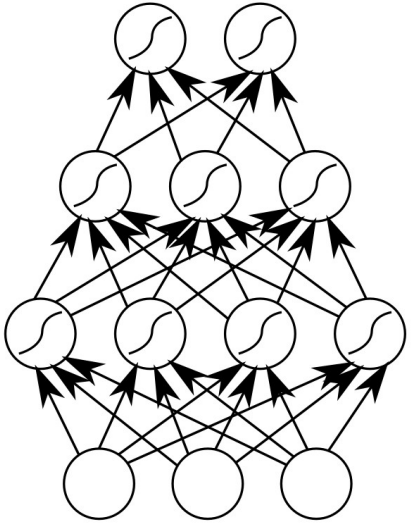
This dissertation is submitted for the degree of
Masters of Philosophy

The big questions

- Where is the frontier of computational creativity?
- How much has deep learning advanced automatic composition?
- How do we evaluate generative models?



vs



Overview

- Sequence modelling for music
  - Motivating example
  - Music primer
  - From Bach Chorales to sequences
- Recurrent neural networks (RNNs)
  - Training and Optimizing BachBot
- Results and the musical Turing test

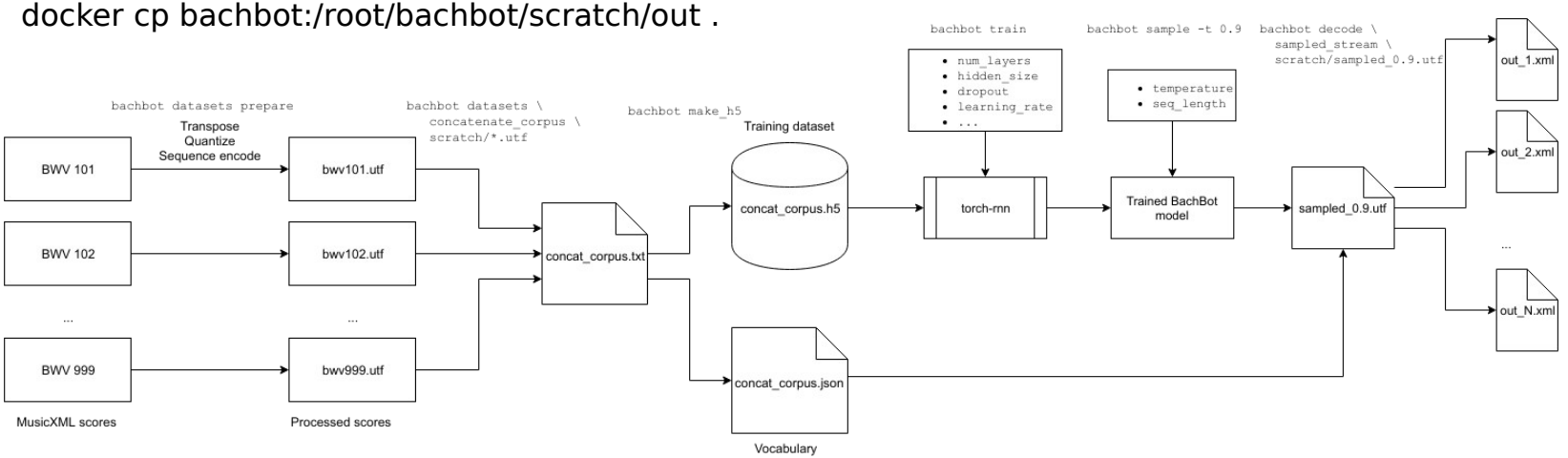
TL;DR

- Deep recurrent neural network model for music capable of:
  - Polyphony
  - Automatic composition
  - Harmonization
- Learns music theory without prior knowledge
- Only 5% out of 1779 participants in a musical Turing test performed better than random chance

For the hands-on type

feynmanliang.github.io/bachbot-slides

```
docker pull fliang/bachbot:aibtb
docker run --name bachbot -it fliang/bachbot:aibtb
bachbot datasets prepare
bachbot datasets concatenate_corpus scratch/BWV-*.utf
bachbot make_h5
bachbot train
bachbot sample ~/bachbot/scratch/checkpoints/*/checkpoint_<ITER>.t7 -t
bachbot decode sampled_stream ~/bachbot/scratch/sampled_$TMP.utf
docker cp bachbot:/root/bachbot/scratch/out .
```



Motivating example for sequence modelling

Sequence modelling of text

The quick brown fox jumps _____

Question: What word comes next?

Sequence modelling of text

The quick brown fox jumps _____

The quick brown fox jumps over
The quick brown fox jumps around
The quick brown fox jumps lackadaisically

Sequence modelling of text

The quick brown fox jumps _____

P(over | the quick brown fox jumps) = 75%
P(around | the quick brown fox jumps) = 24%
P(lackadaisically | the quick brown fox jumps) = 1%

Sequence modelling of text

The quick brown fox jumps _____

P(over | the quick brown fox jumps) = 75%
P(around | the quick brown fox jumps) = 24%
P(lackadaisically | the quick brown fox jumps) = 1%

Question: Any potential problems?

The 2-gram sequence model

The quick brown fox jumps _____

P(over | jumps) = 90%
P(around | jumps) = 8%
P(behind | jumps) = 2%

Generating text using the 2-gram sequence model:

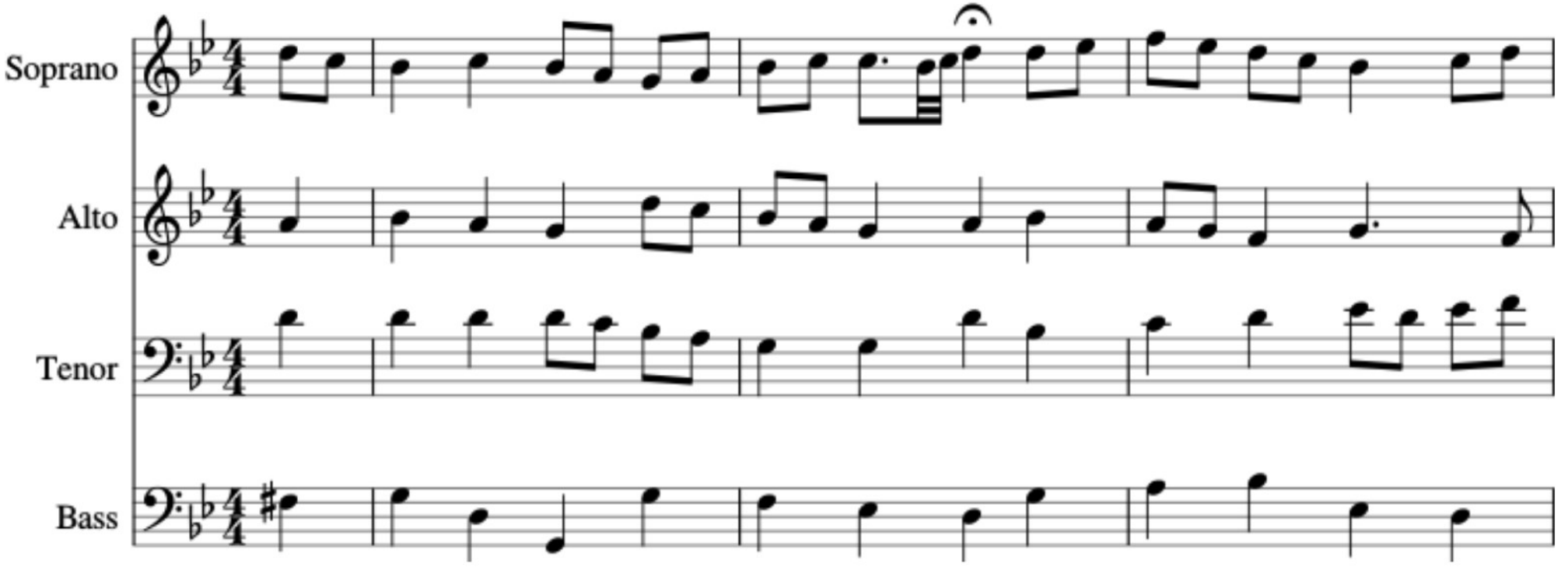
I
I am
am enjoying
enjoying GOTO
GOTO Berlin
I am enjoying GOTO Berlin

n-gram models trained on Hamlet

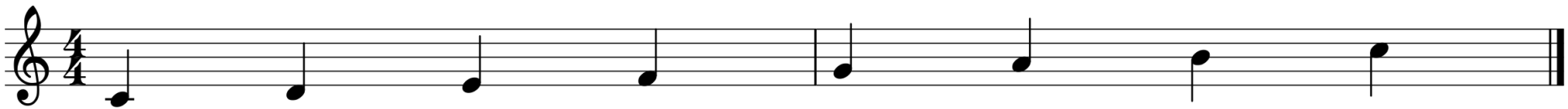
- 1-gram
  - To him swallowed confess hear both. Which. Of save on trail for are ay device androte life have
- 2-gram
  - Why dost stand forth thy canopy, forsooth; he is this palpable hit the King Henry. Liveking. Follow.
- 4-gram
  - King Henry. What! I will go seek the traitor Gloucester. Exeunt some of the watch. Agreat banquet serv’d in;

Examples taken from Dan Jurafsky, https://web.stanford.edu/~jurafsky/slp3/slides/LM_4.pdf

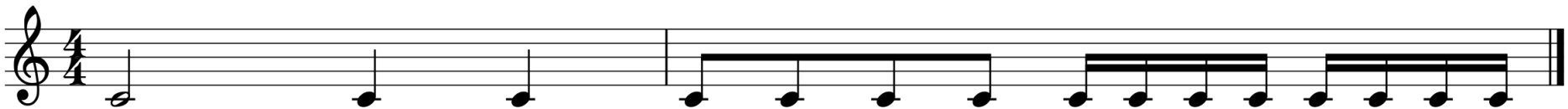
Music Primer



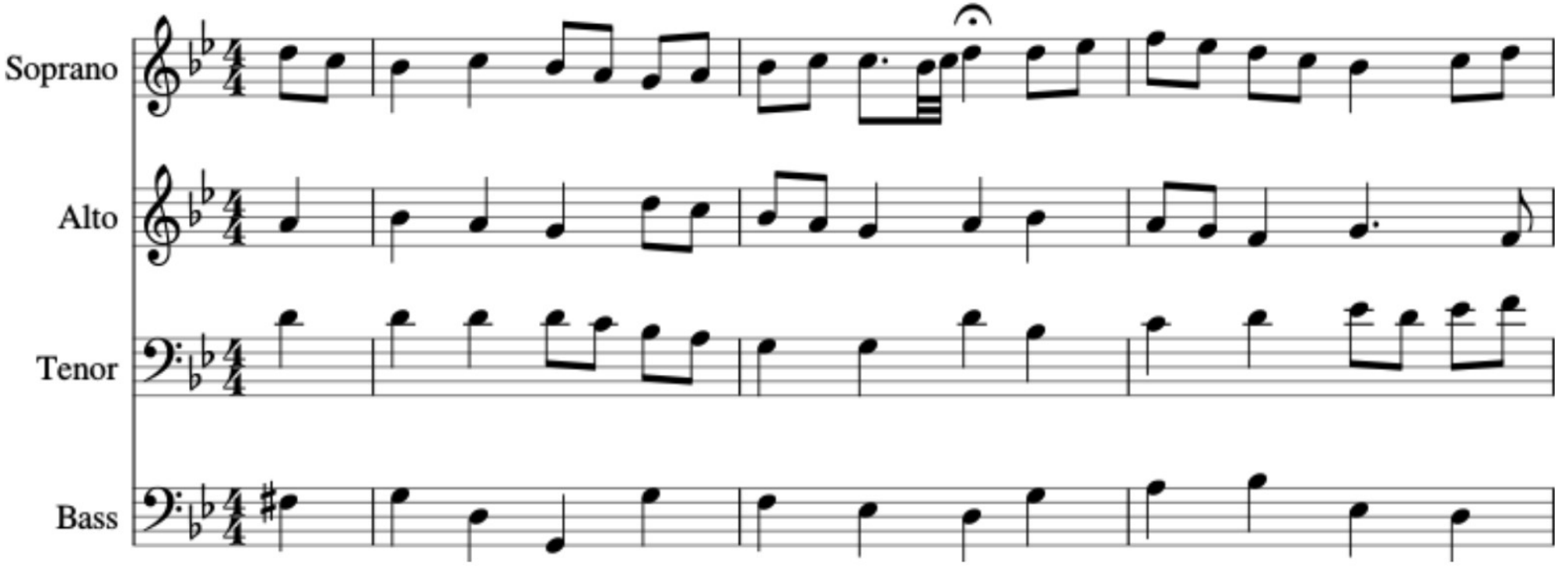
Modern music notation



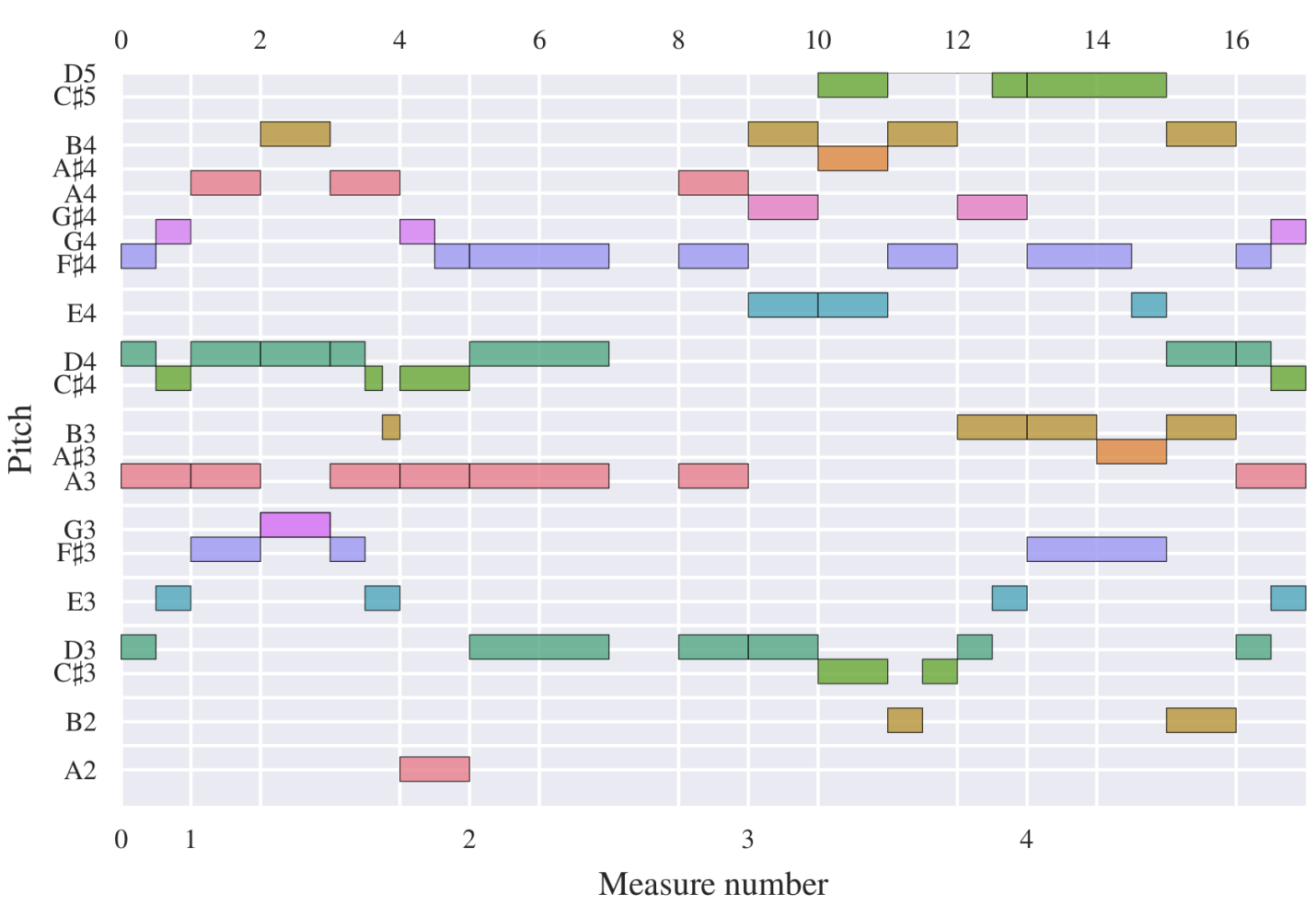
Pitch: how “high” or “low” a note is



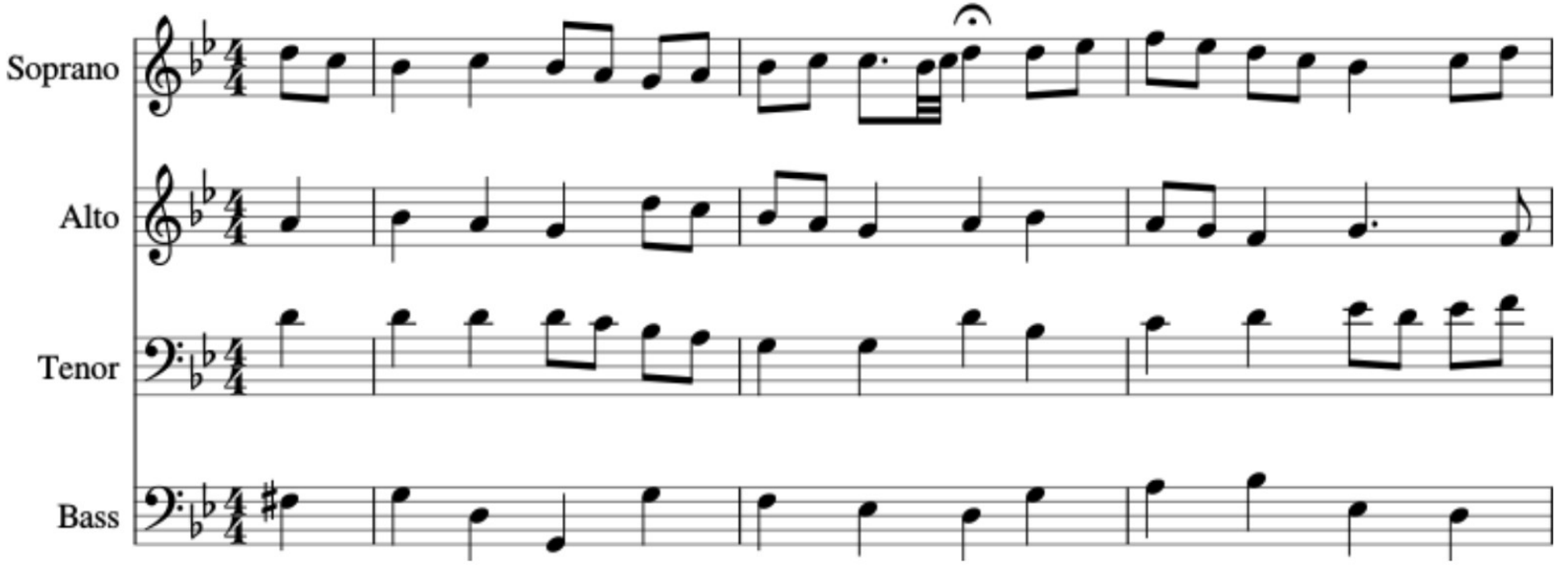
Duration: how “long” a note is



Polyphony: multiple simultaneous voices

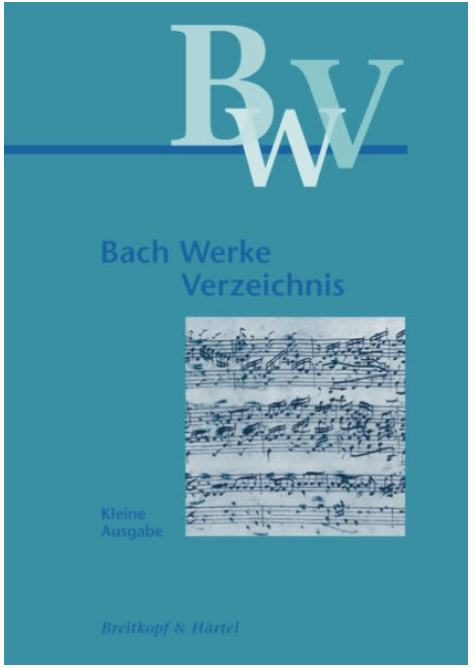


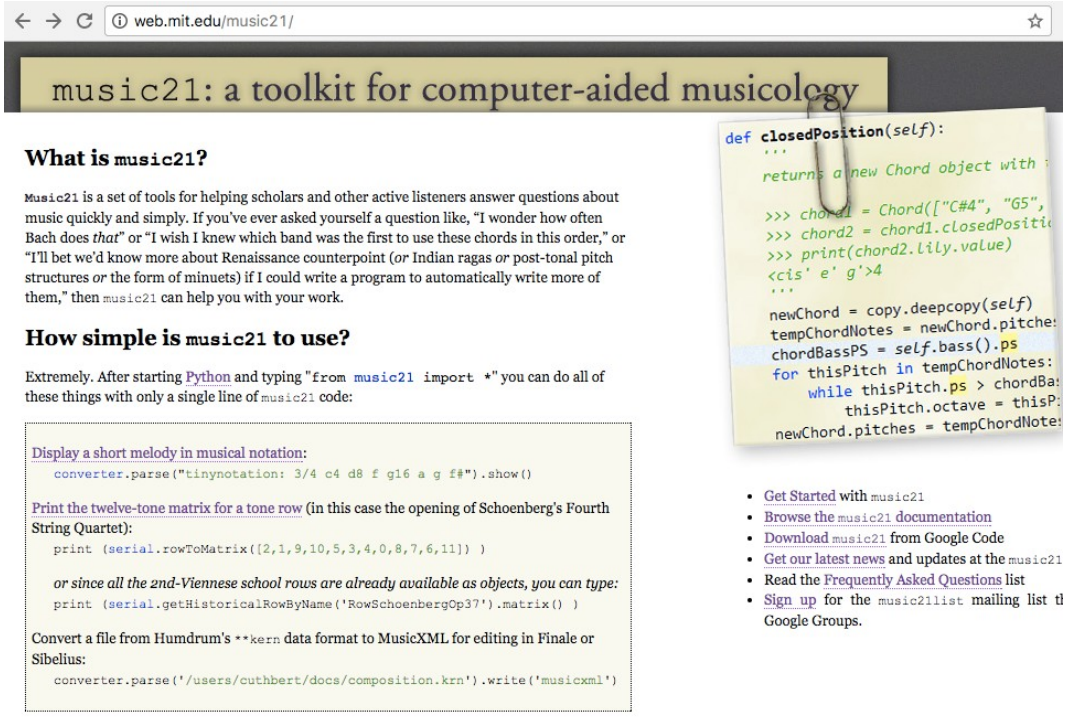
Piano roll: convenient computational representation



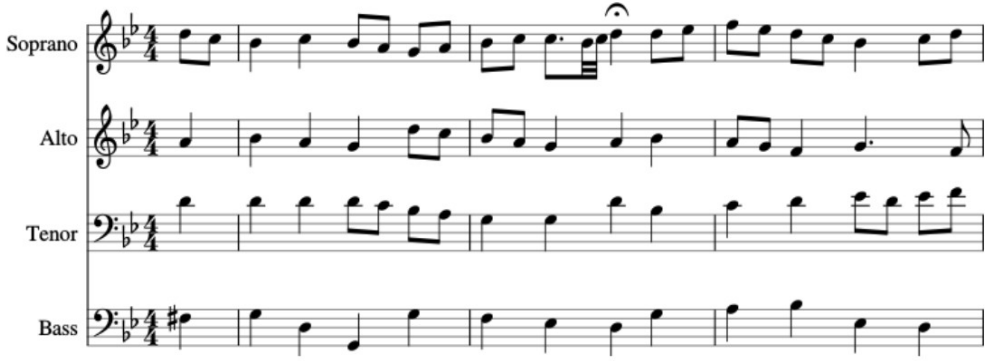
Fermatas and **phrasing**

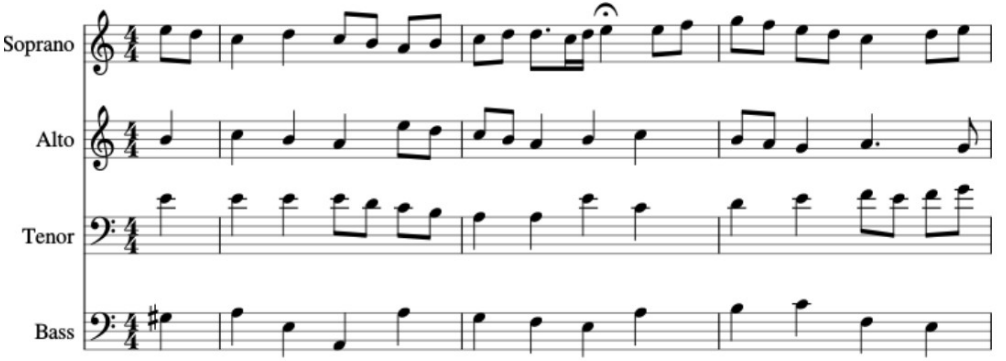
From Bach Chorales to Sequences



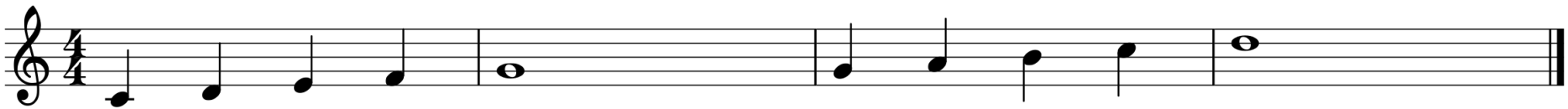


http://web.mit.edu/music21/

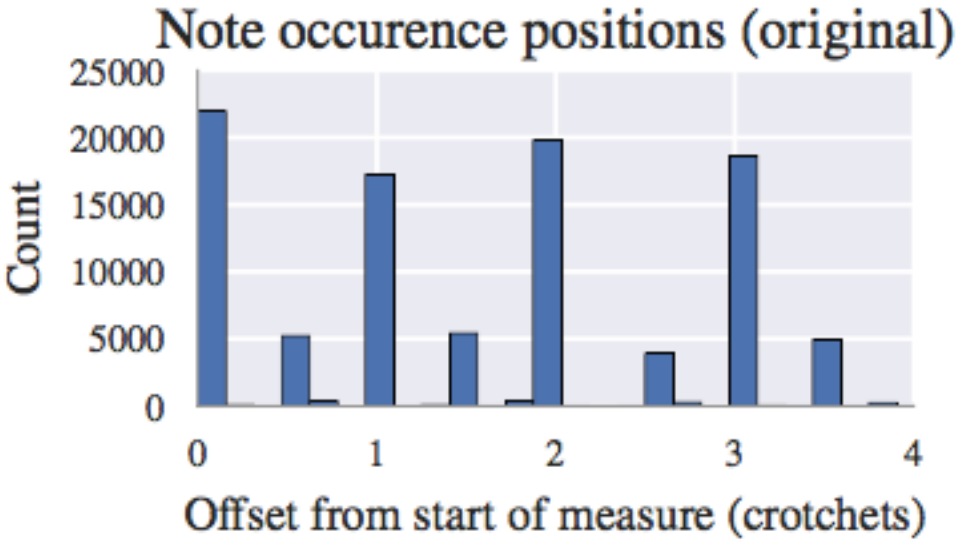


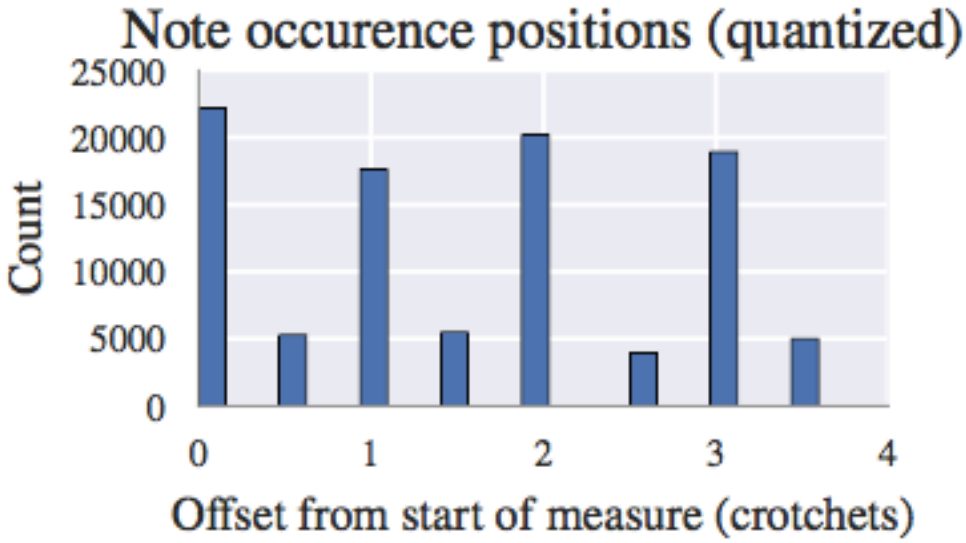


Transpose to Cmaj/Amin (convenience)
Quantize to 16th notes (computational)



Transposition preserves relative pitches





Quantization to 16th notes: affects less than 0.2% of dataset

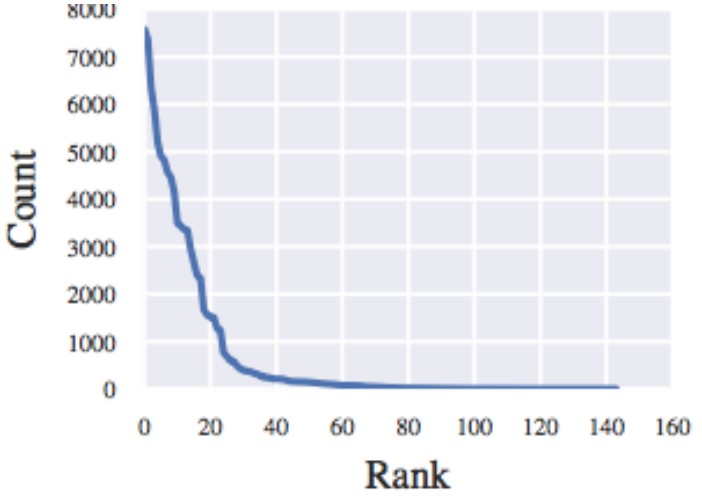
Handling polyphony

Question: How many chords can be constructed from 4 voices, each with 128 pitches?

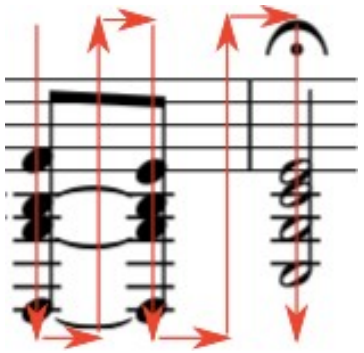
Handling polyphony

Question: How many chords can be constructed from 4 voices, each with 128 pitches?

Answer: $O(128^4)$! Data sparsity issue…



Serialize in SATB order

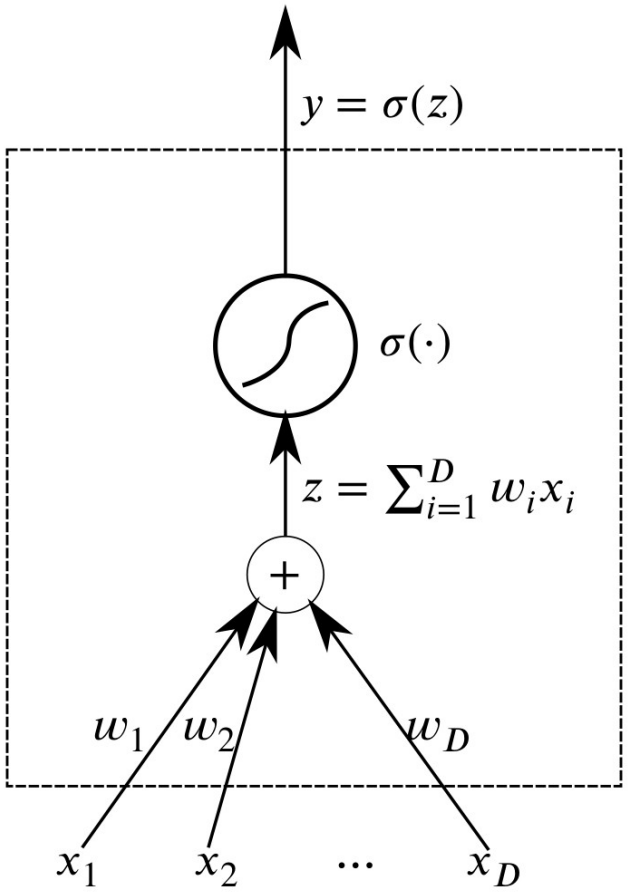


START
(62, True)
(56, True)
(52, True)
(47, True)
|||
(60, False)
(56, True)
(52, True)
(47, True)
|||

(.)
(60, False)
(57, False)
(52, False)
(50, False)
|||
(.)
(60, True)
(57, True)
(52, True)
(50, True)
|||

$O(128^4) \Rightarrow O(128)$ vocab. size!

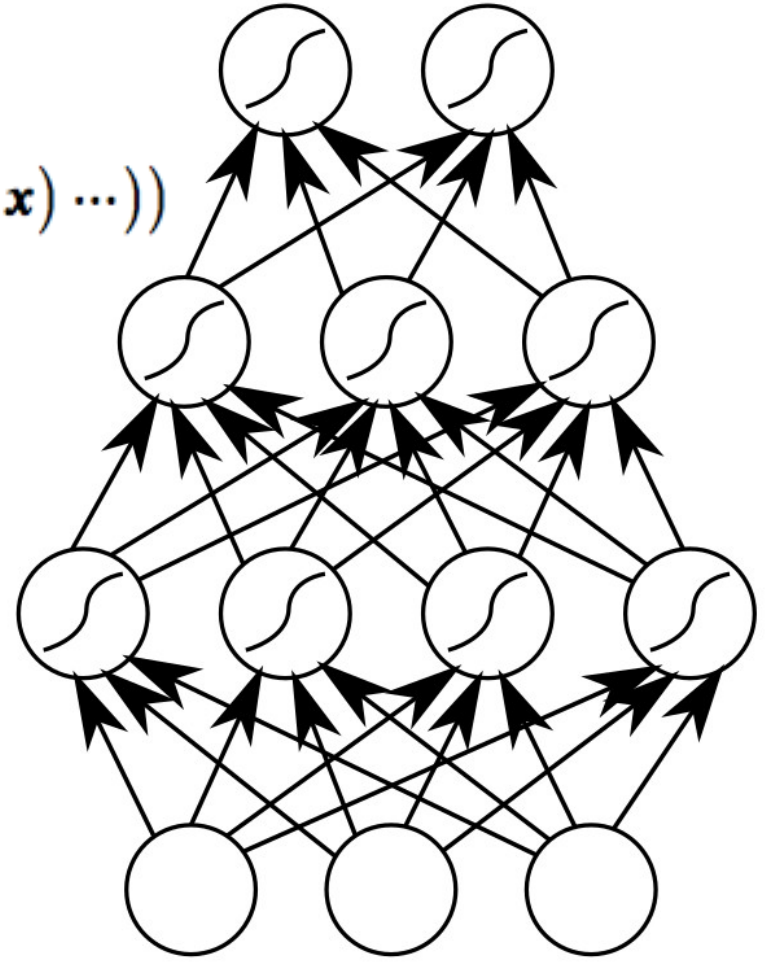
Recurrent neural networks (RNNs)



Neuron

Input **x**, output **y**, parameters **w**, activations **z**

$$\sigma\left(\boldsymbol{W}_{L,L-1}\sigma\left(\boldsymbol{W}_{L-1,L-2}\cdots\sigma\left(\boldsymbol{W}_{2,1}\boldsymbol{x}\right)\cdots\right)\right)$$



Output Layer

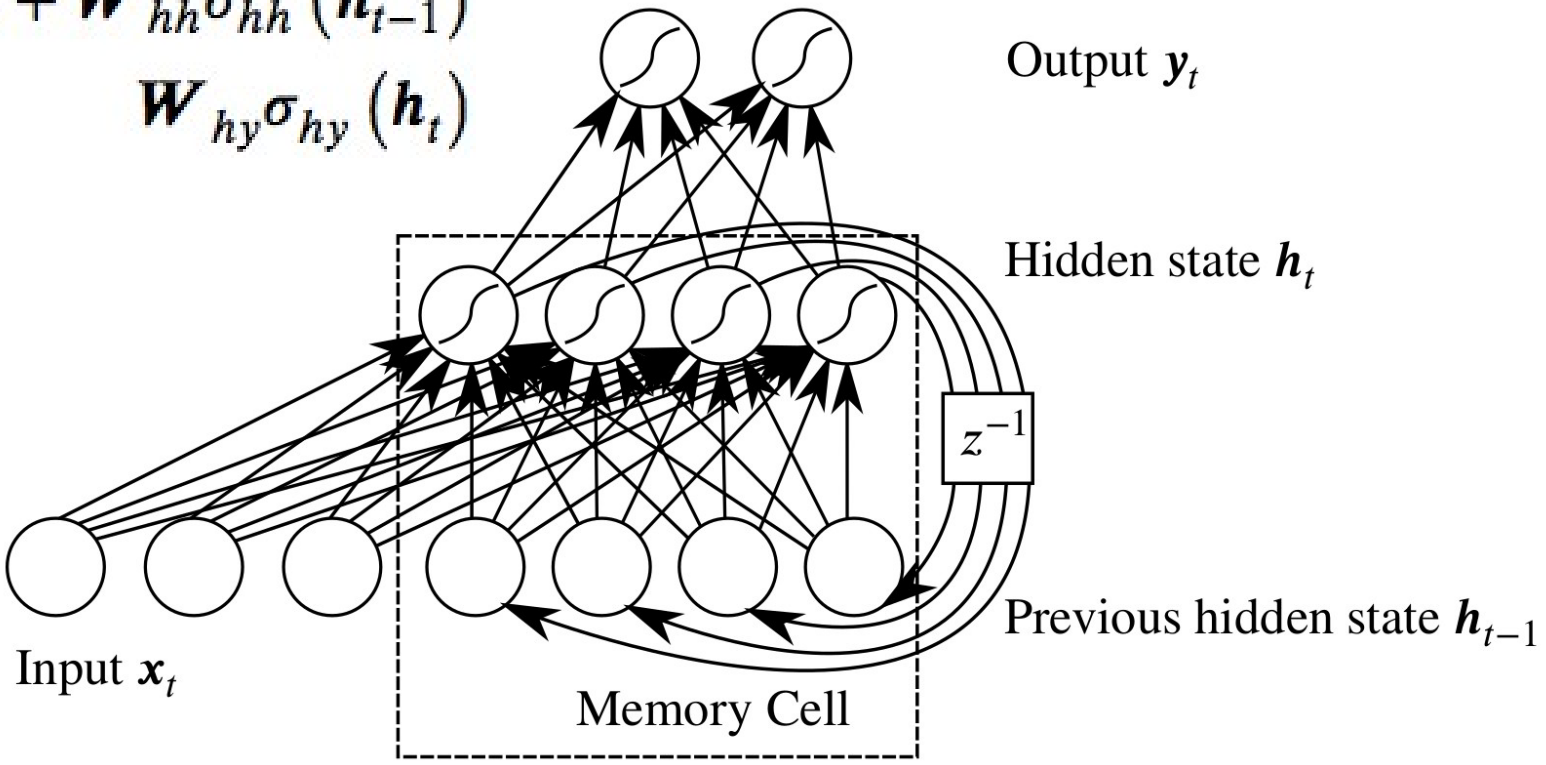
Hidden Layers

Input Layer

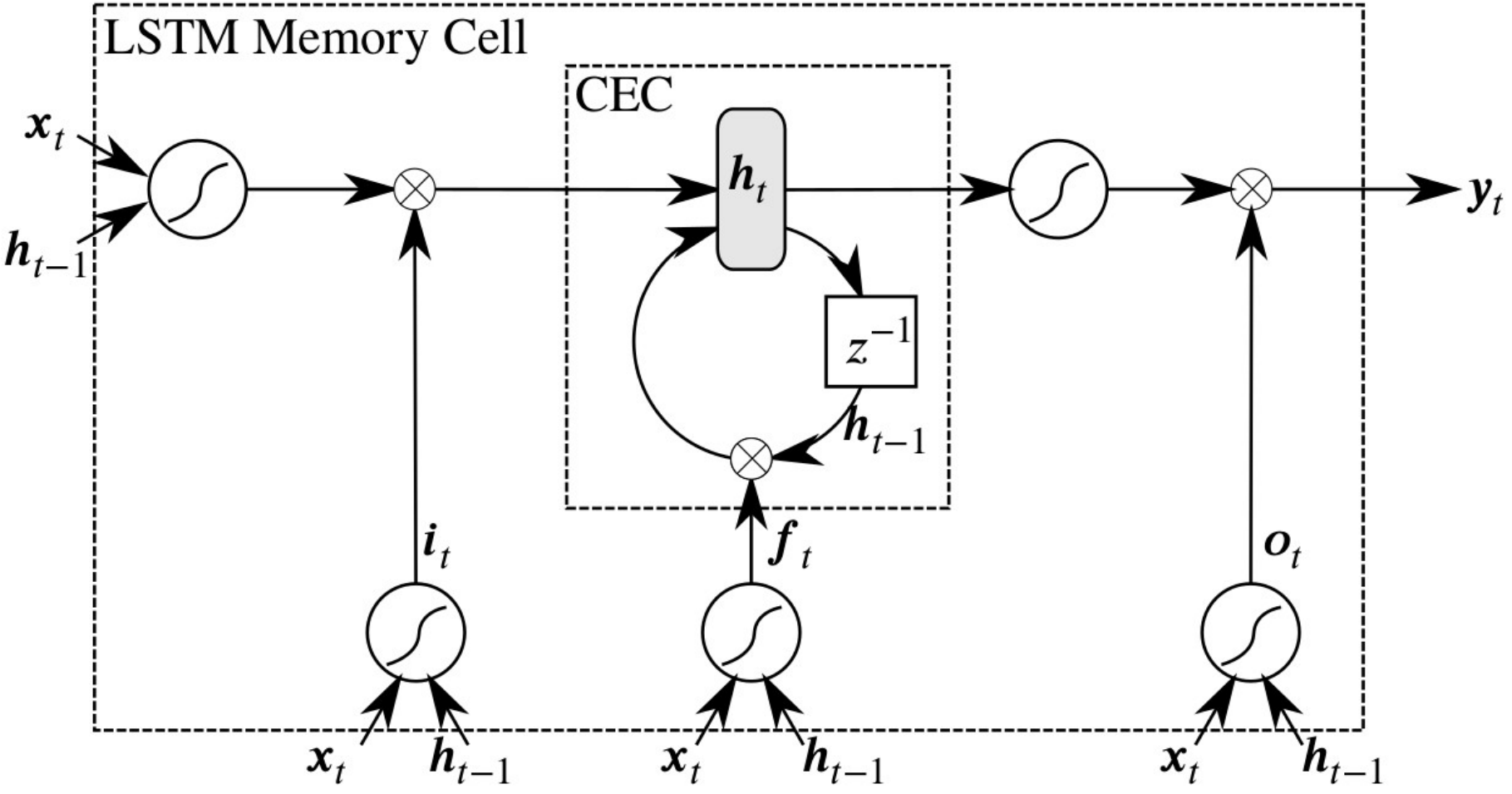
Feedforward neural network

$$\boldsymbol{h}_t = \boldsymbol{W}_{xh}\sigma_{xh}\left(\boldsymbol{x}_t\right) + \boldsymbol{W}_{hh}\sigma_{hh}\left(\boldsymbol{h}_{t-1}\right)$$

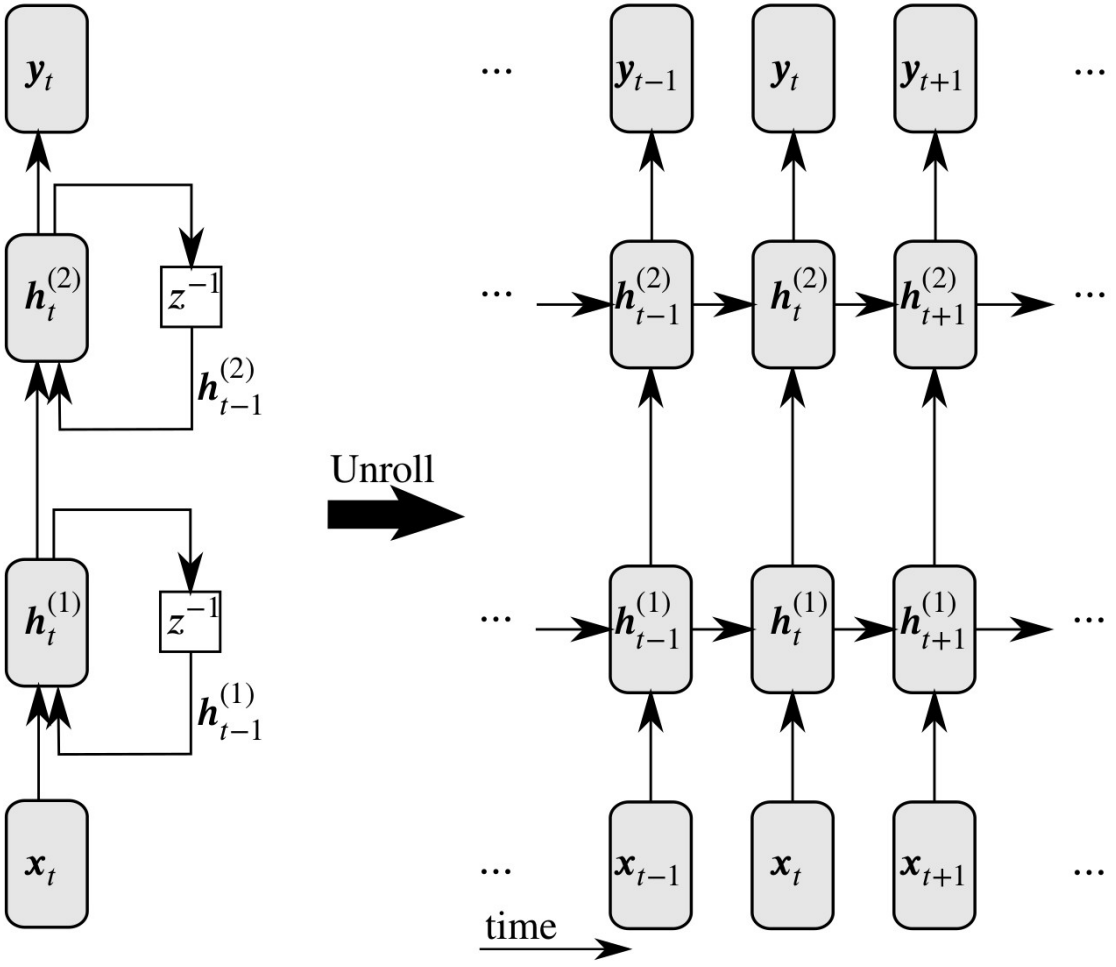
$$\boldsymbol{y}_t = \boldsymbol{W}_{hy}\sigma_{hy}\left(\boldsymbol{h}_t\right)$$



Memory cell

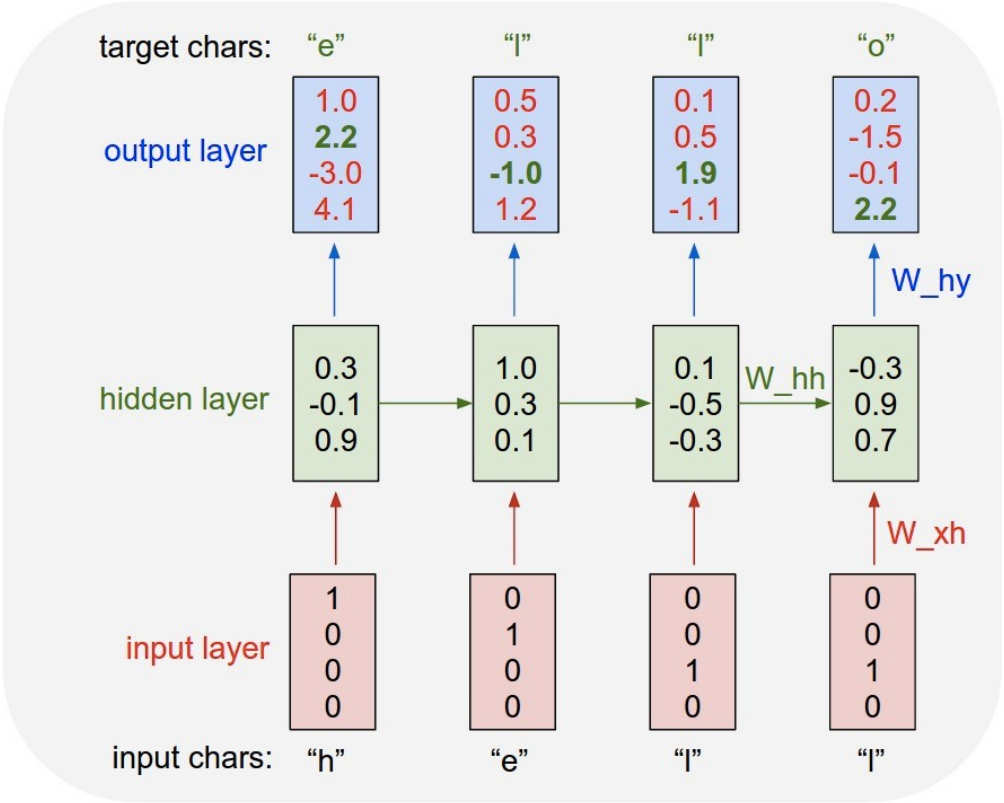


Long short-term memory (**LSTM**) cell



Stacking memory cells to form a deep RNN
Unrolling for training

Sequential prediction training criteria



https://karpathy.github.io/2015/05/21/rnn-effectiveness/

Training and optimizing BachBot

Training BachBot

RNN dynamics

$$\tilde{P}(x_{t+1}|x_t, h_{t-1})$$

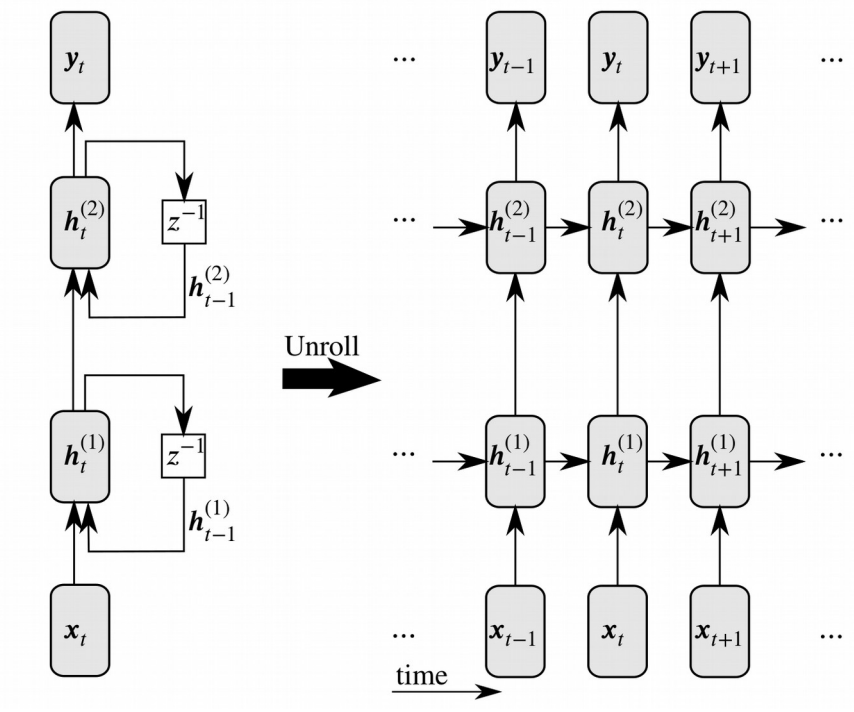
Initial state (all 0s)

$$+ \quad h_0$$

Prob. distr. over sequences

$$y_t = \tilde{P}(x_t|x_{1:t-1})$$

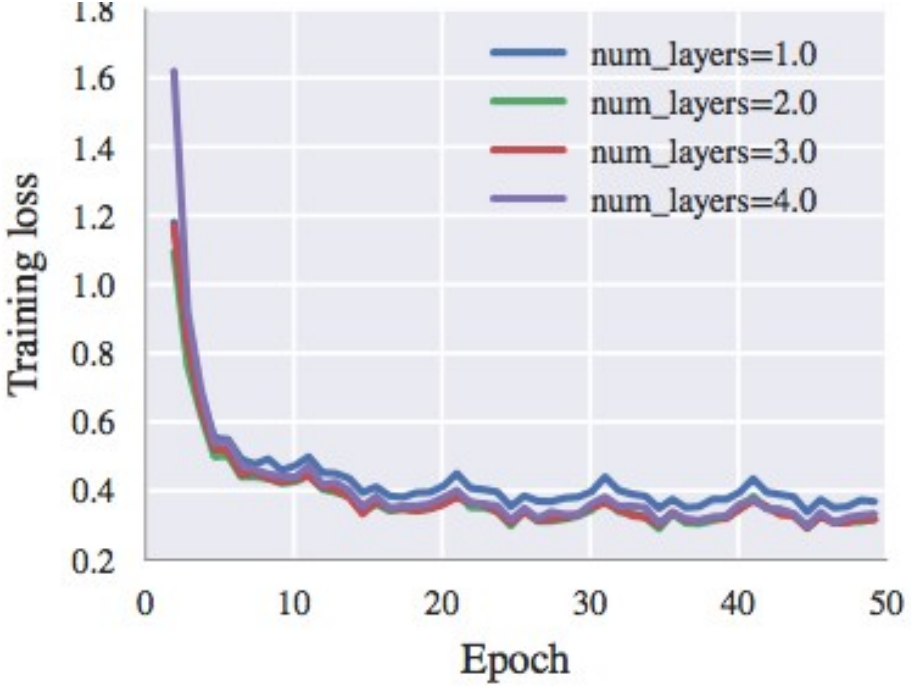
Optimize the RNN parameters...

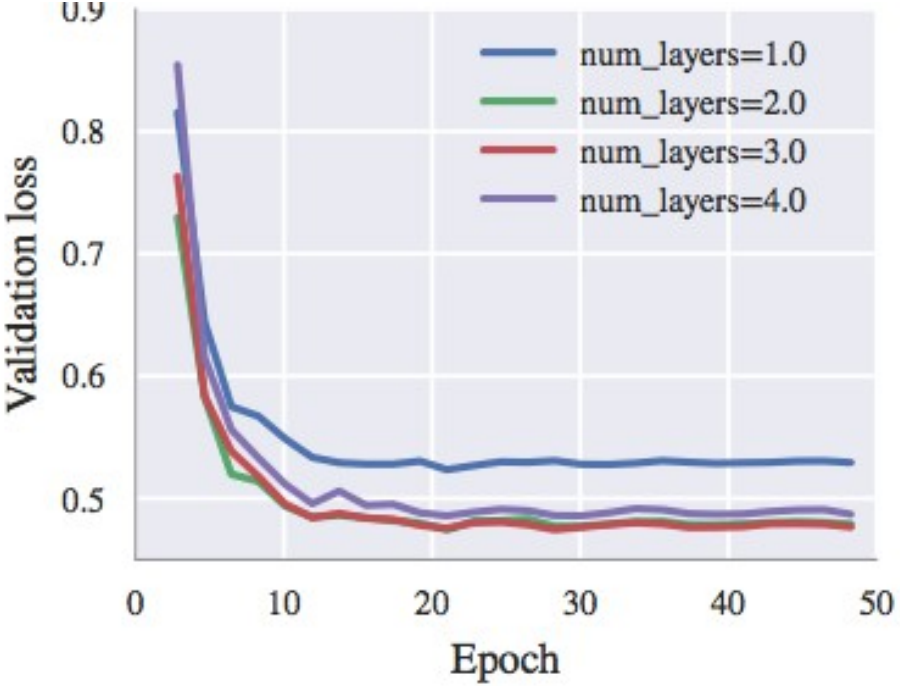


...in order to maximize the probability of the real Bach chorales.

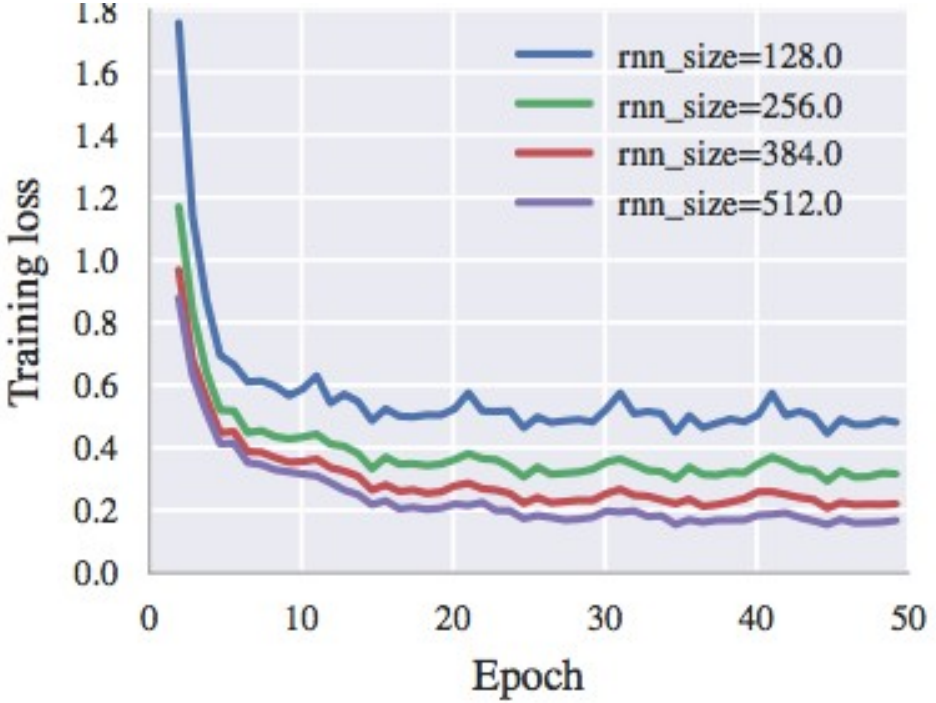
	Single Batch		30 Epochs (seconds)
	mean (sec)	std (sec)	(minutes)
CPU	4.287	0.311	256.8
GPU	0.513	0.001	28.5

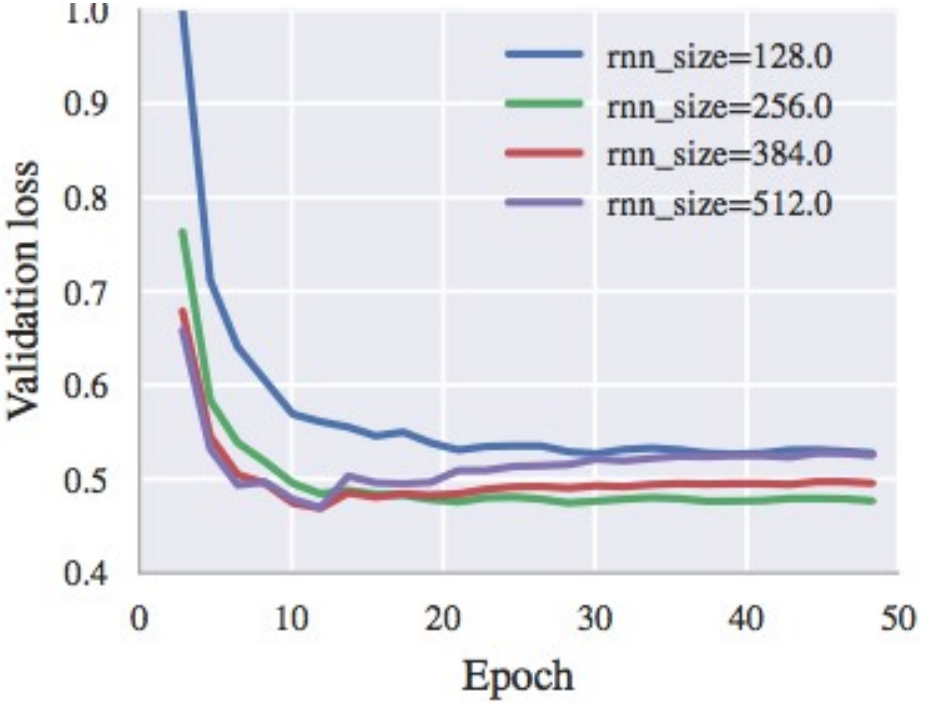
GPUs deliver a 8x performance speedup



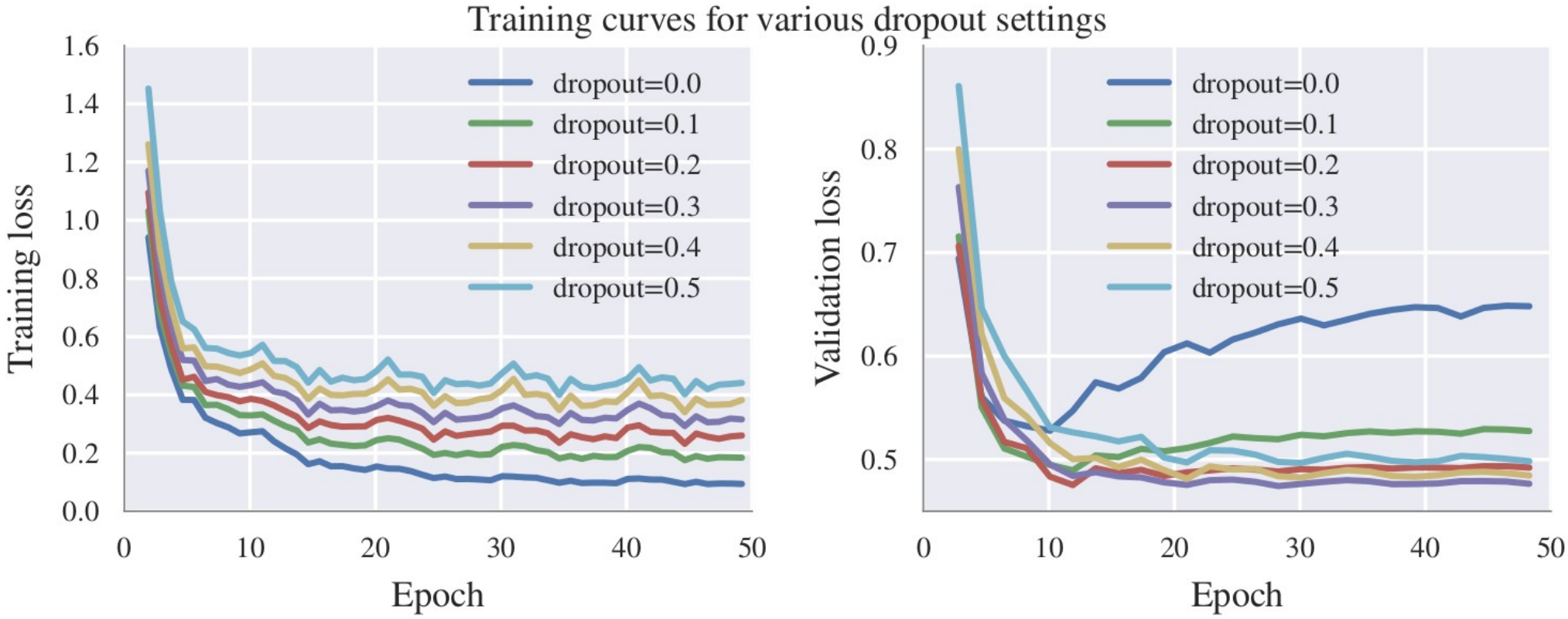


Depth matters (to a certain point)

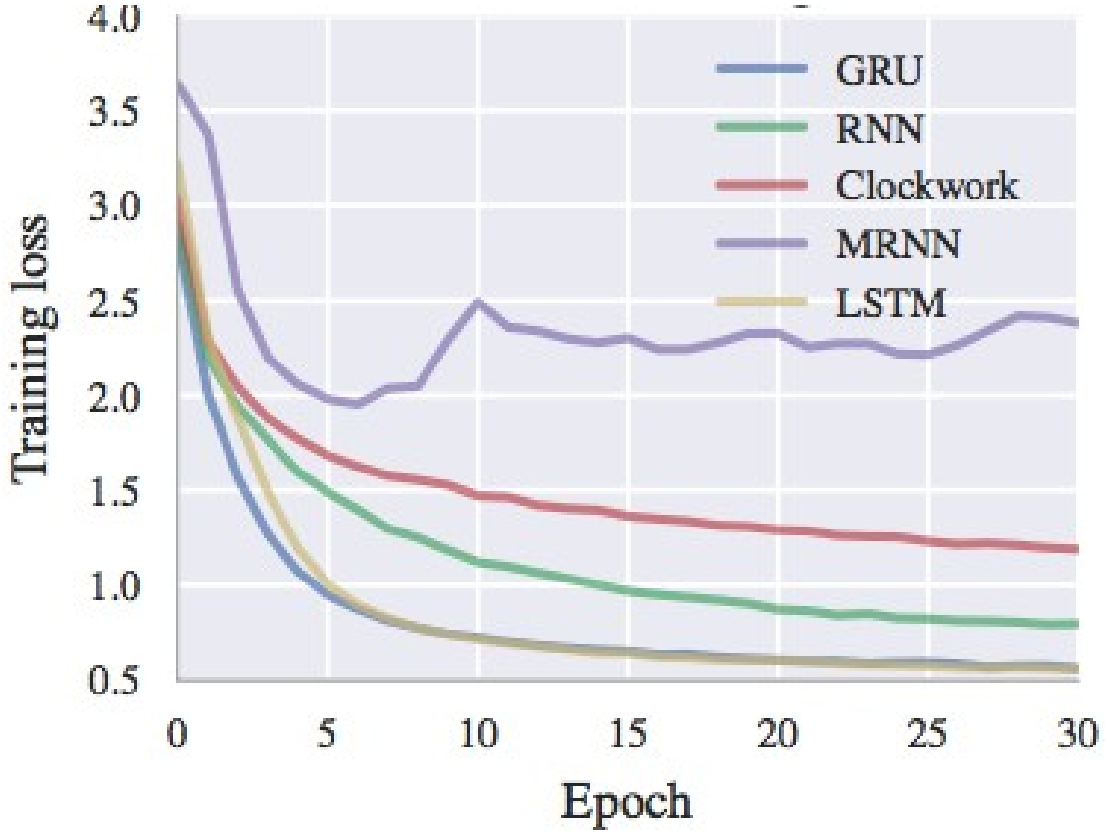




Hidden state size matters (to a certain point)



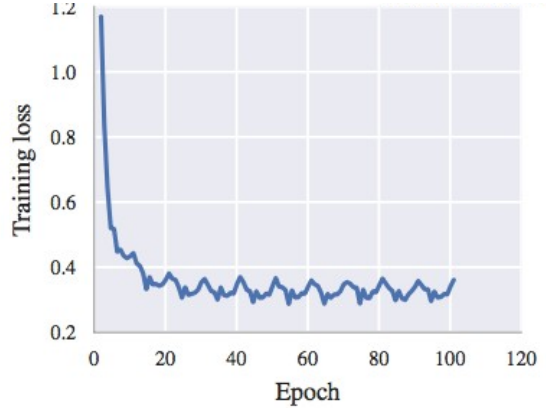
Dropout improves generalization (to a certain point)

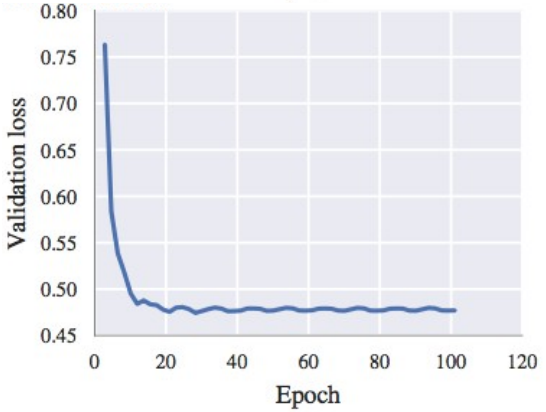


LSTM memory cells matter

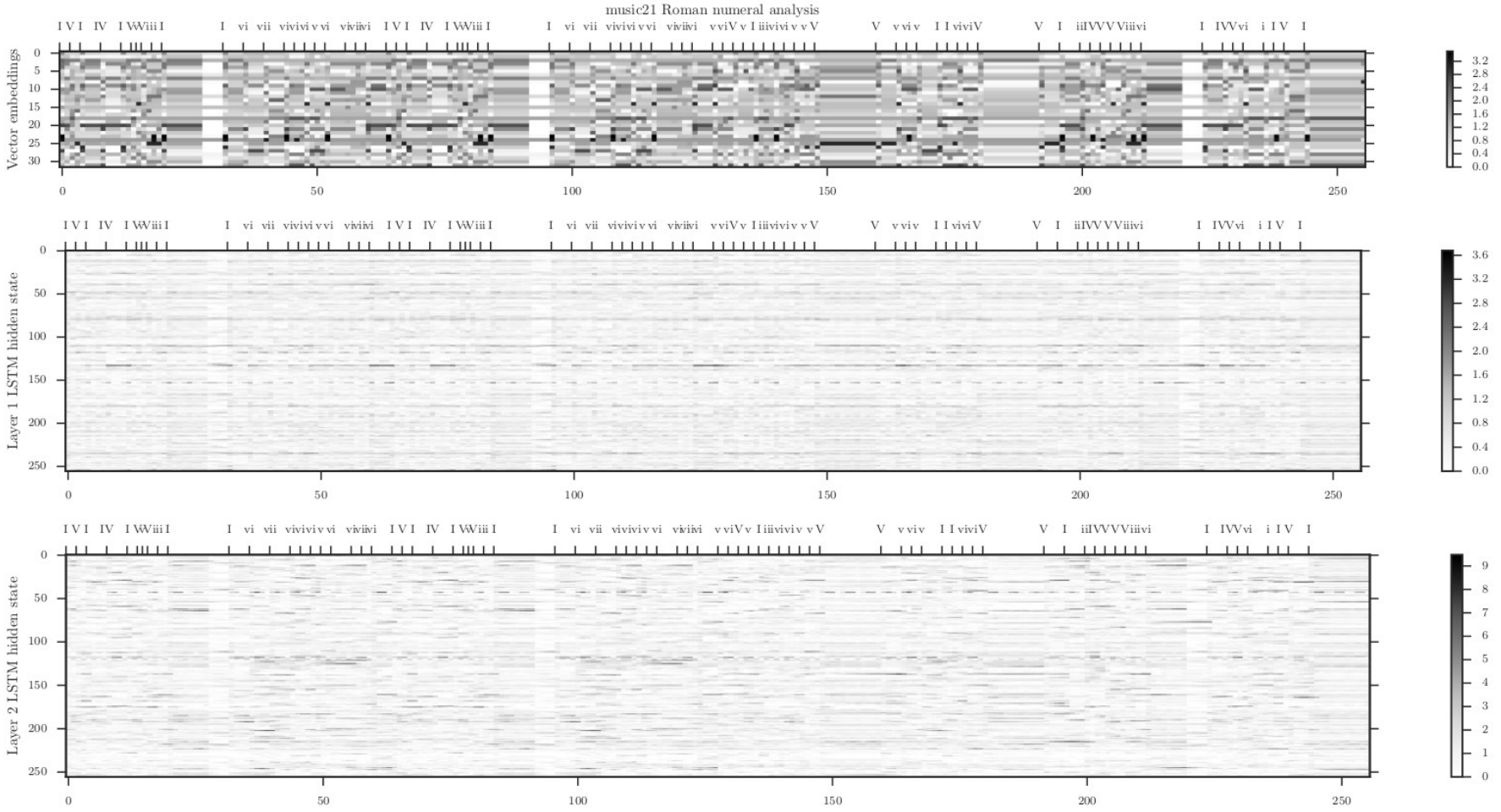
The final model

- Architecture
  - 32-dimensional vector space embedding
  - 3 layered, 256 hidden unit LSTMs
- Standard deep learning tricks
  - 30% dropout
  - Batch normalization
  - 128 timestep truncated back-propagation through time (BPTT)
- Bach chorales specific tricks
  - SATB order: S determines harmonic context
  - Explicitly encode fermatas



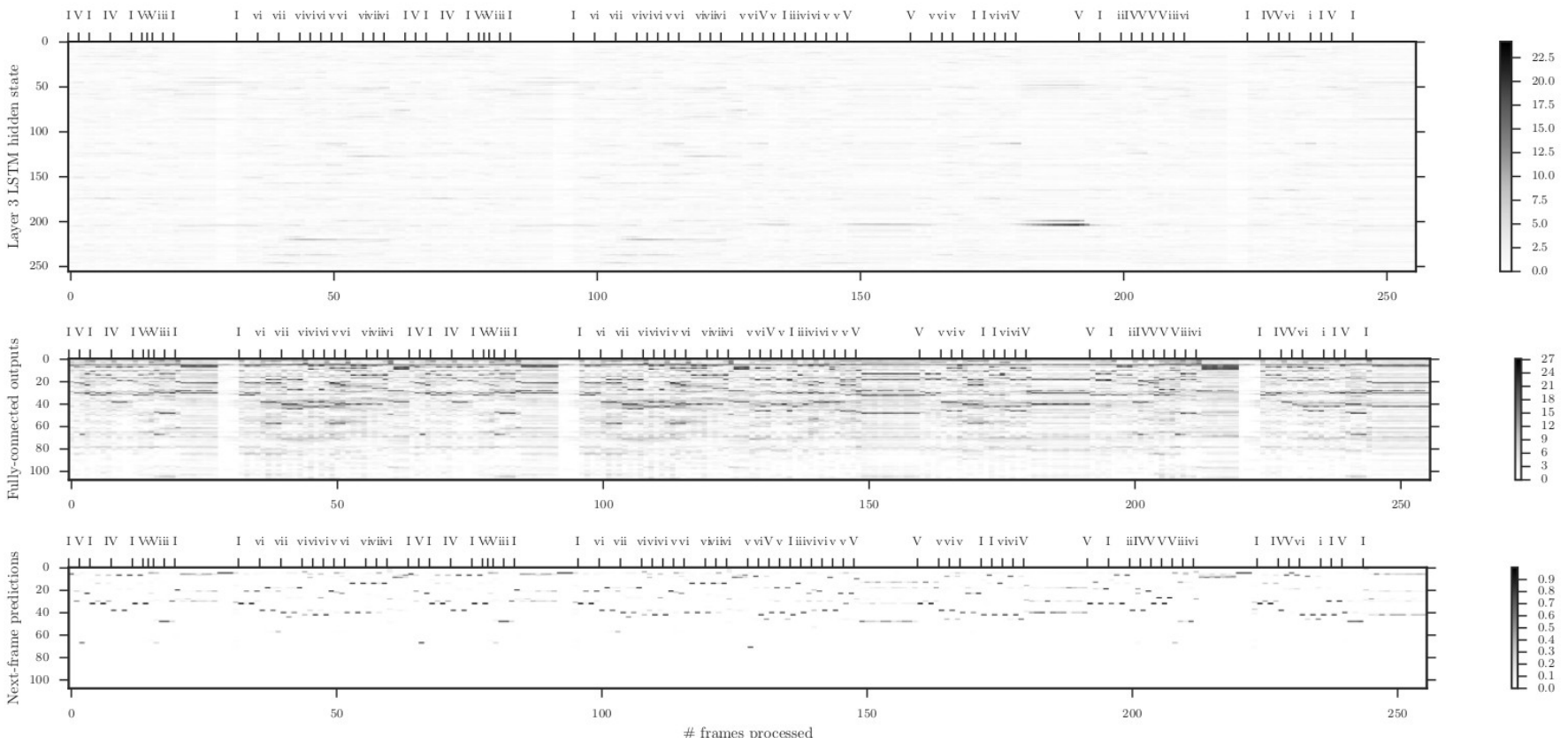


Results



Hidden state is difficult to interpret!

Input and memory cell (layer 1 and 2)

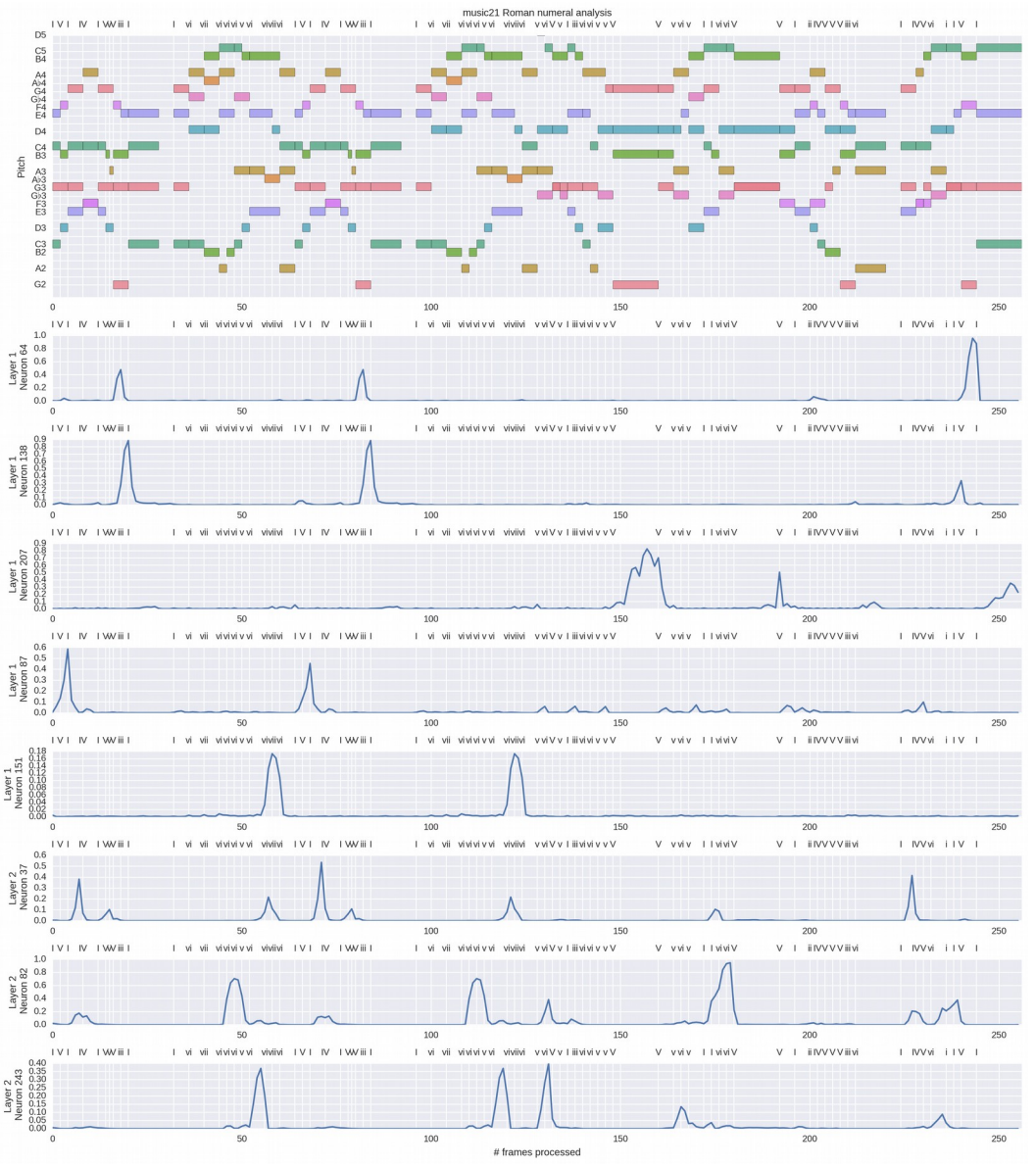


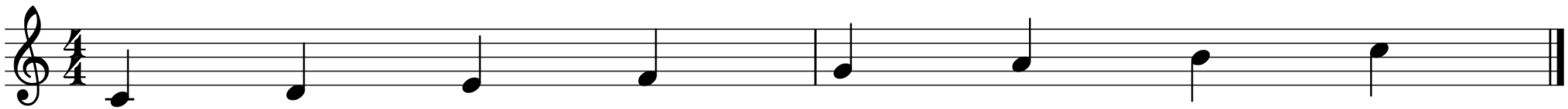
Layers closer to the output resemble piano roll (consequence of sequential training criteria)

Memory cell (layer 3), outputs, and predictions

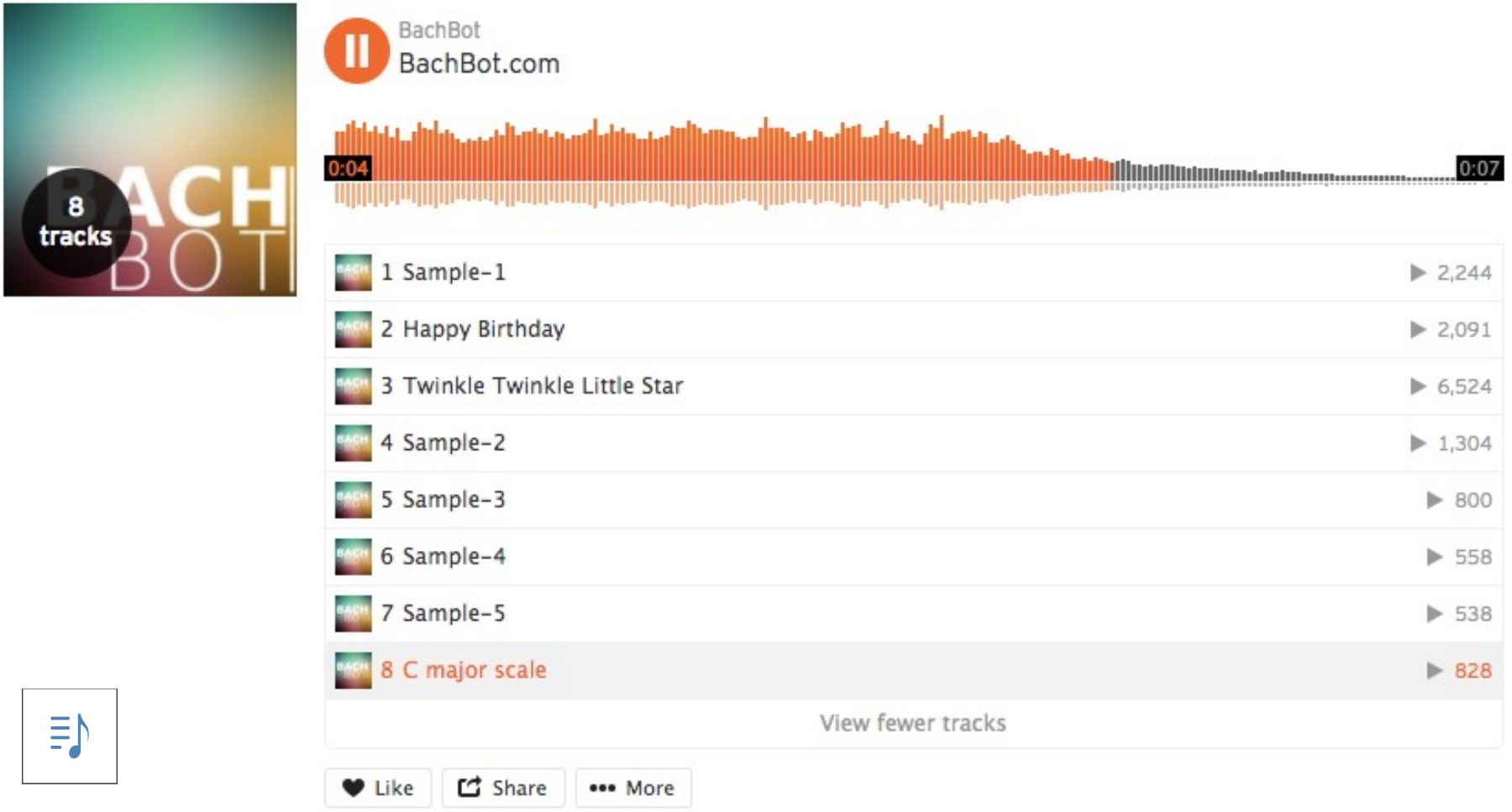
Model learns music theory!

- L1N64 and L1N138: Perfect cadences with root position chords in tonic key
- L1N151: A minor cadences ending phrases 2 and 4
- L1N87 and L2N37: I^6 chords





Harmonization: given a melody (here C major scale)



Harmonization: produce the accompanying parts

https://soundcloud.com/bachbot

BachBot
BachBot.com

0:03 0:27

3 Twinkle Twinkle Little Star	▶ 6,524
4 Sample-2	▶ 1,304
5 Sample-3	▶ 800
6 Sample-4	▶ 558
7 Sample-5	▶ 538

View 8 tracks

Like Share ••• More

What if Bach remixed twinkle twinkle little star?

https://soundcloud.com/bachbot

Musical Turing test

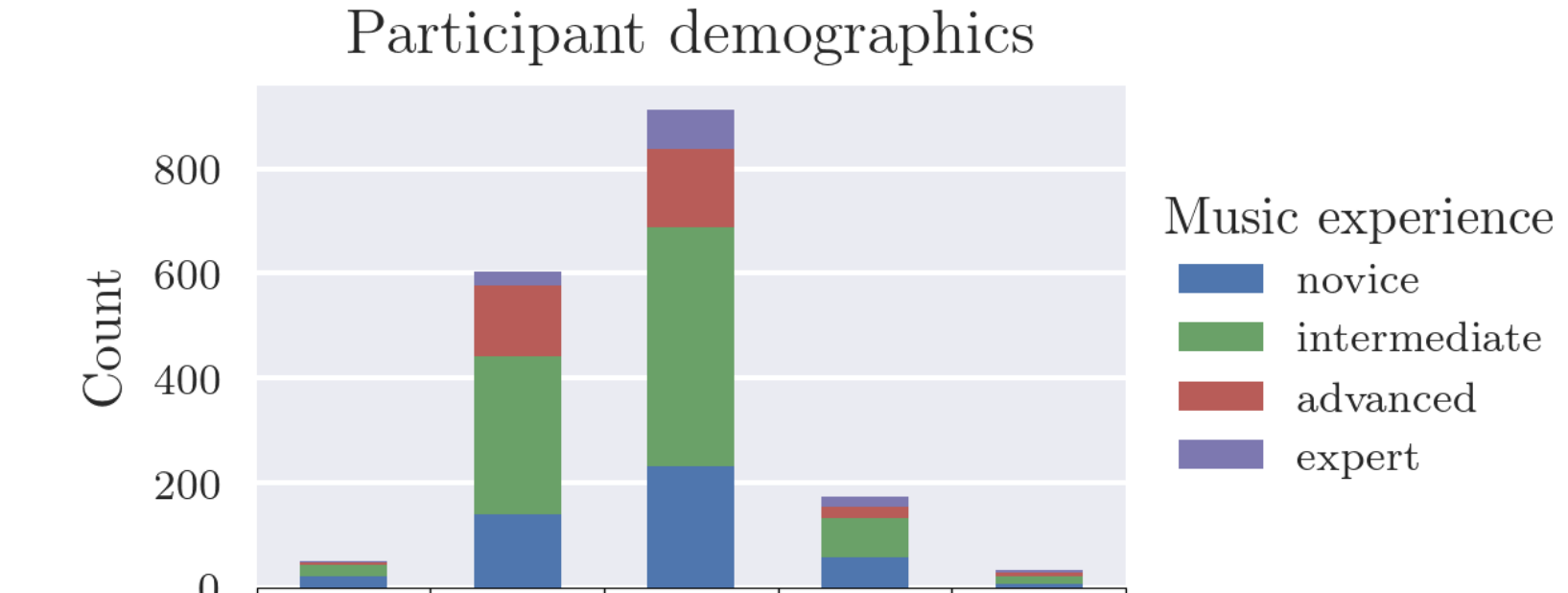
http://bachbot.com

The BachBot Challenge

	Select the music most similar to Bach
Select	
Select	
	Submit

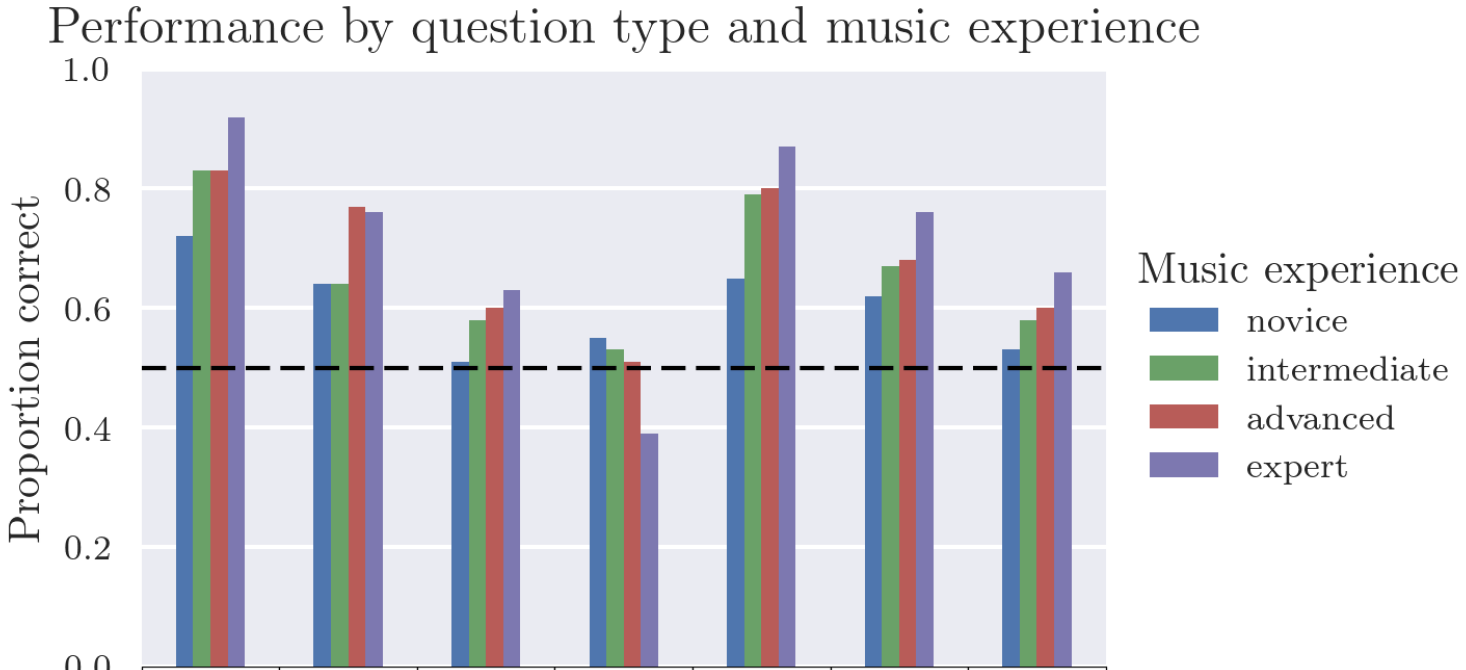
20%

Question 1 out of 5



	under18	18to25	26to45	46to60	over60
novice	23	141	233	57	8
intermediate	20	302	454	75	14
advanced	6	134	151	23	7
expert	3	27	75	20	6

Participants by age group and music experience



	S	A	T	B	AT	ATB	SATB
novice	0.72	0.64	0.51	0.55	0.65	0.62	0.53
intermediate	0.83	0.64	0.58	0.53	0.79	0.67	0.58
advanced	0.83	0.77	0.6	0.51	0.8	0.68	0.6
expert	0.92	0.76	0.63	0.39	0.87	0.76	0.66

More experienced respondents tend to do better

Closing Remarks

- Deep LSTM model for composing, completing, and generating Bach Chorales
- Open source (github.com/feynmanliang/bachbot), integrated into Google Magenta's "Polyphonic RNN" model (magenta.tensorflow.org)
- Appears to learn music theoretic concepts without prior knowledge
- Largest music Turing test to date with over 1779 participants, average participant performs **only 5% better than random guessing**

Thank You!

Please

Remember to rate this session

Thank you!

@GOTOber

gotober.com