

Please

Ask questions
through the app



Rate Session

Thank you!



Let's talk about Bach



BachBot
BachBot.com



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

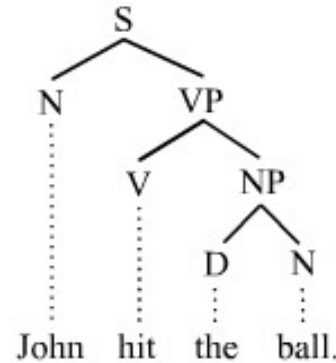
Department of Engineering
University of Cambridge

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

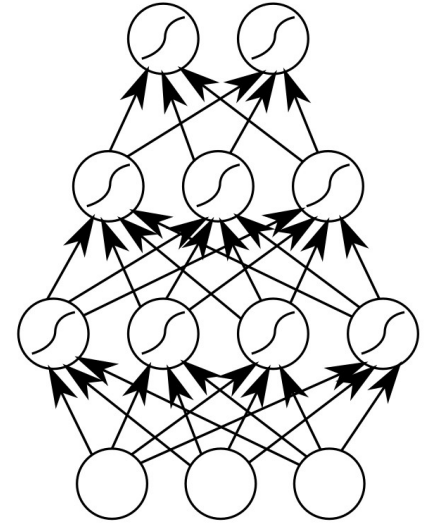
This dissertation is submitted for the degree of
Masters of Philosophy

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

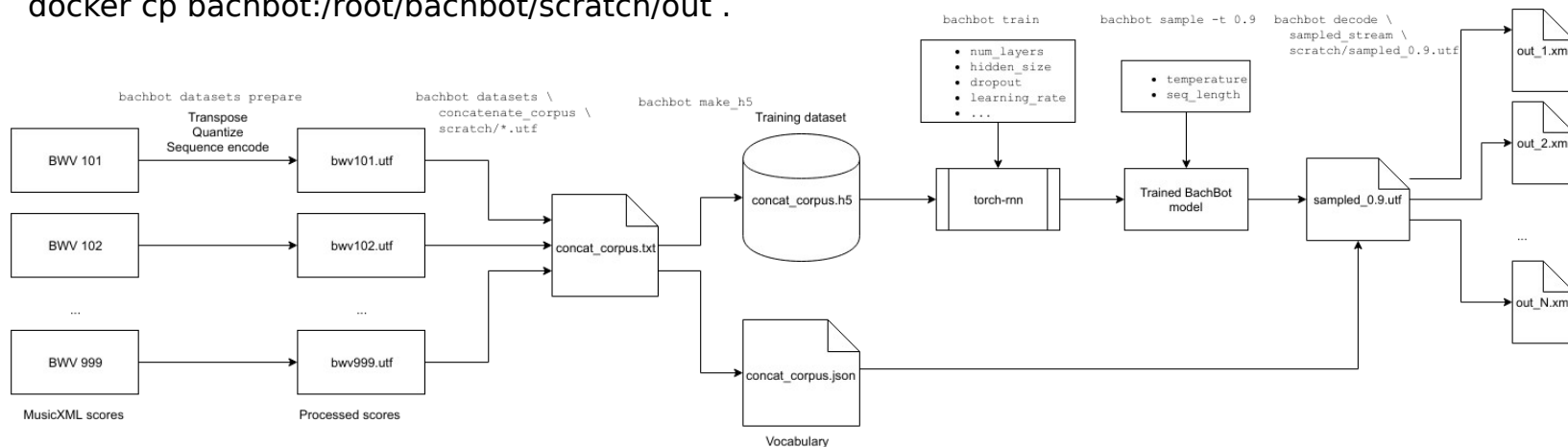
TL;DR

- 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_$.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 _____

$P(\text{over} \mid \text{the quick brown fox jumps}) = 75\%$

$P(\text{around} \mid \text{the quick brown fox jumps}) = 24\%$

$P(\text{lackadaisically} \mid \text{the quick brown fox jumps}) = 1\%$

Question: Any potential problems?

The 2-gram sequence model

The quick brown fox jumps _____

$P(\text{over} \mid \text{jumps}) = 90\%$

$P(\text{around} \mid \text{jumps}) = 8\%$

$P(\text{behind} \mid \text{jumps}) = 2\%$

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. A great banquet serv'd in;

Music Primer

Soprano

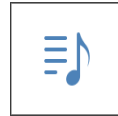
Alto

Tenor

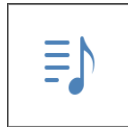
Bass

A musical score for four voices: Soprano, Alto, Tenor, and Bass. The score is written in 4/4 time and the key of B-flat major (two flats). The Soprano part uses a treble clef and features a melodic line with a fermata on the eighth measure. The Alto part also uses a treble clef and follows a similar melodic contour. The Tenor part uses a bass clef and provides a harmonic accompaniment with eighth and sixteenth notes. The Bass part uses a bass clef and features a more active line with a sharp sign on the first measure, possibly indicating a key signature change or a specific pitch.

Modern music notation



Pitch: how “high” or “low” a note is



Duration: how “long” a note is

Soprano

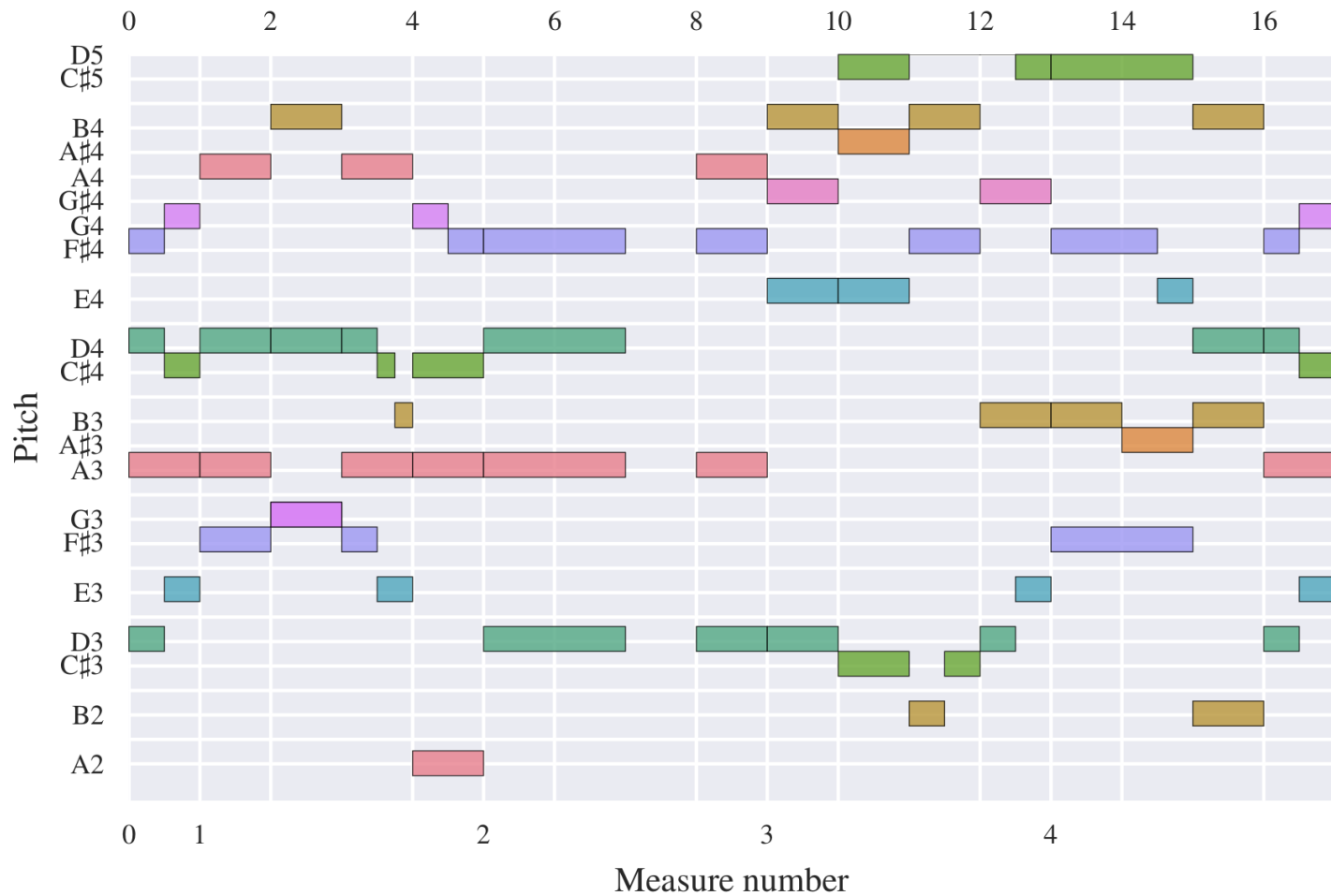
Alto

Tenor

Bass

The image displays a musical score for four voices: Soprano, Alto, Tenor, and Bass. The key signature is B-flat major (two flats) and the time signature is 4/4. The Soprano part begins with a half note B-flat, followed by quarter notes C, D, E, F, G, A, B-flat, and C. The Alto part begins with a half note B-flat, followed by quarter notes C, D, E, F, G, A, B-flat, and C. The Tenor part begins with a half note B-flat, followed by quarter notes C, D, E, F, G, A, B-flat, and C. The Bass part begins with a half note B-flat, followed by quarter notes C, D, E, F, G, A, B-flat, and C. The Soprano part features a melodic line with a half note B-flat and a quarter note C in the final measure. The Alto part features a melodic line with a half note B-flat and a quarter note C in the final measure. The Tenor part features a melodic line with a half note B-flat and a quarter note C in the final measure. The Bass part features a melodic line with a half note B-flat and a quarter note C in the final measure.

Polyphony: multiple simultaneous voices



Piano roll: convenient computational representation

Soprano

Alto

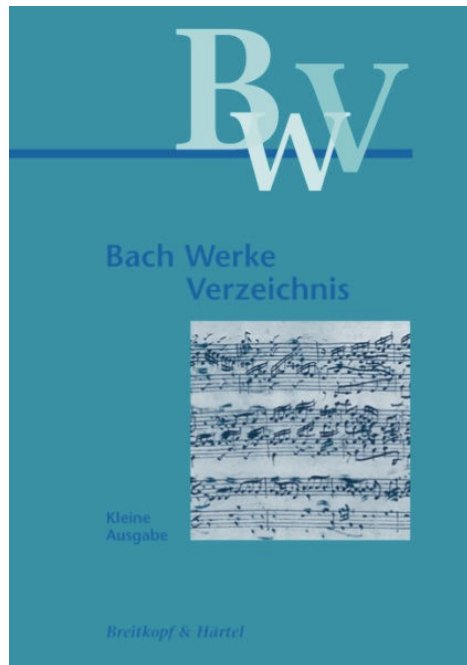
Tenor

Bass

The image displays a musical score for four voices: Soprano, Alto, Tenor, and Bass. The key signature is B-flat major (two flats) and the time signature is 4/4. The Soprano part begins with a treble clef and a key signature of two flats. It features a fermata on the final note of the third measure. The Alto part also uses a treble clef and a key signature of two flats. The Tenor part uses a bass clef and a key signature of two flats. The Bass part uses a bass clef and a key signature of two flats, with a sharp sign on the first line of the staff. The score is written in a standard musical notation style with notes, rests, and a fermata.

Fermatas and phrasing

From Bach Chorales to Sequences



← → ↻ ⓘ web.mit.edu/music21/ ☆

music21: a toolkit for computer-aided musicology

What is music21?

Music21 is a set of tools for helping scholars and other active listeners answer questions about music quickly and simply. If you've ever asked yourself a question like, "I wonder how often Bach does *that*" or "I wish I knew which band was the first to use these chords in this order," or "I'll bet we'd know more about Renaissance counterpoint (or Indian ragas or post-tonal pitch structures or the form of minuets) if I could write a program to automatically write more of them," then music21 can help you with your work.

How simple is music21 to use?

Extremely. After starting Python and typing "from music21 import *" you can do all of these things with only a single line of music21 code:

```
Display a short melody in musical notation:
converter.parse("tinynotation: 3/4 c4 d8 f g16 a g f#").show()

Print the twelve-tone matrix for a tone row (in this case the opening of Schoenberg's Fourth String Quartet):
print (serial.rowToMatrix([2,1,9,10,5,3,4,0,8,7,6,11]) )

or since all the 2nd-Viennese school rows are already available as objects, you can type:
print (serial.getHistoricalRowByName('RowSchoenbergOp37').matrix() )

Convert a file from Humdrum's **kern data format to MusicXML for editing in Finale or Sibelius:
converter.parse('/users/cuthbert/docs/composition.krn').write('musicxml')
```

```
def closedPosition(self):
    """
    returns a new Chord object with :

    >>> chord1 = Chord(["C#4", "G5",
    >>> chord2 = chord1.closedPosition
    >>> print(chord2.lily.value)
    <cis' e' g'>4
    """
    newChord = copy.deepcopy(self)
    tempChordNotes = newChord.pitches
    chordBassPS = self.bass().ps
    for thisPitch in tempChordNotes:
        while thisPitch.ps > chordBassPS:
            thisPitch.octave = thisPitch.octave + 1
    newChord.pitches = tempChordNotes
```

- [Get Started with music21](#)
- [Browse the music21 documentation](#)
- [Download music21](#) from Google Code
- [Get our latest news](#) and updates at the music21
- [Read the Frequently Asked Questions](#) list
- [Sign up](#) for the music21list mailing list or Google Groups.

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

Soprano

Alto

Tenor

Bass

4/4

Soprano

Alto

Tenor

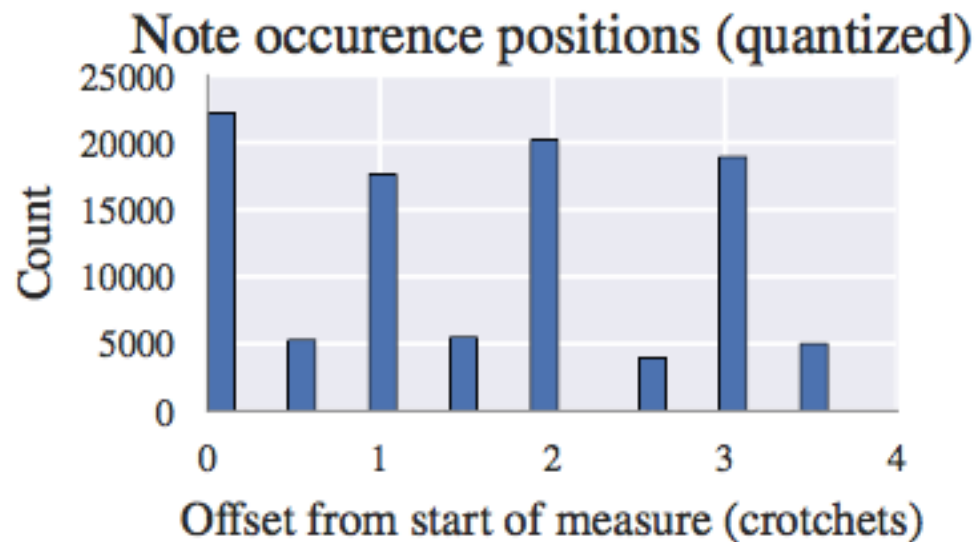
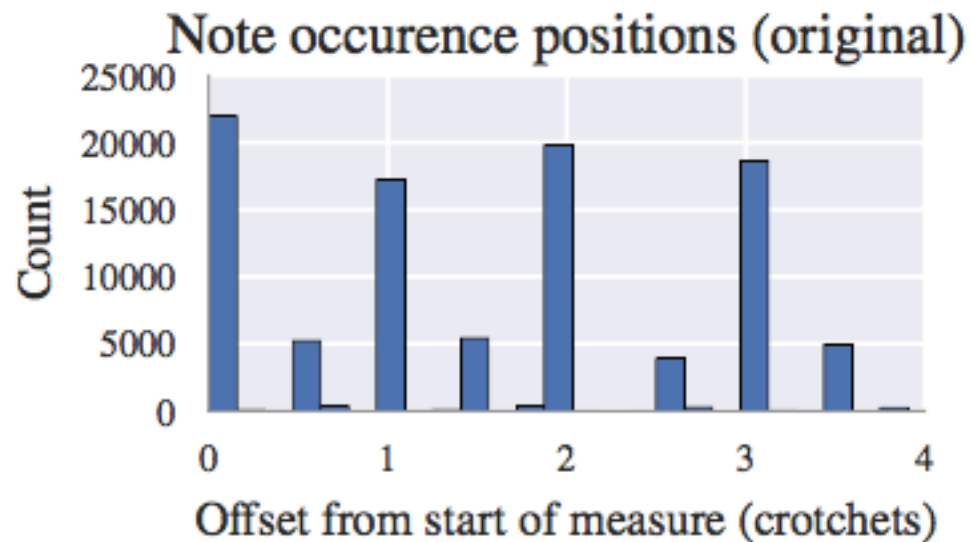
Bass

4/4

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

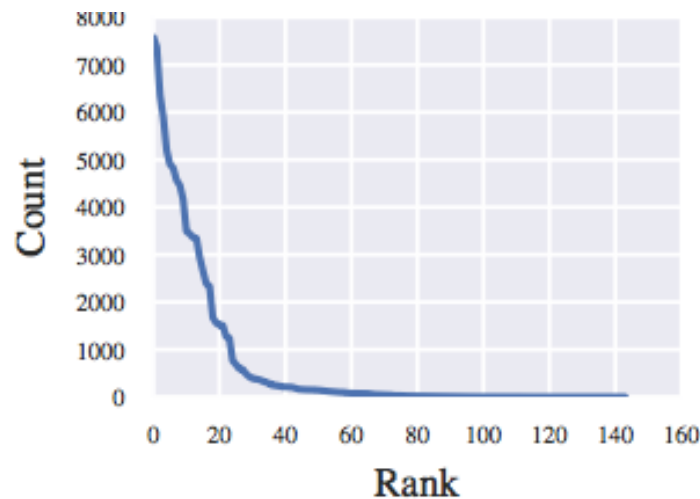
Handling polyphony

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

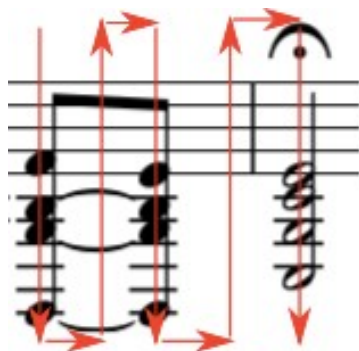
Handling polyphony

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

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



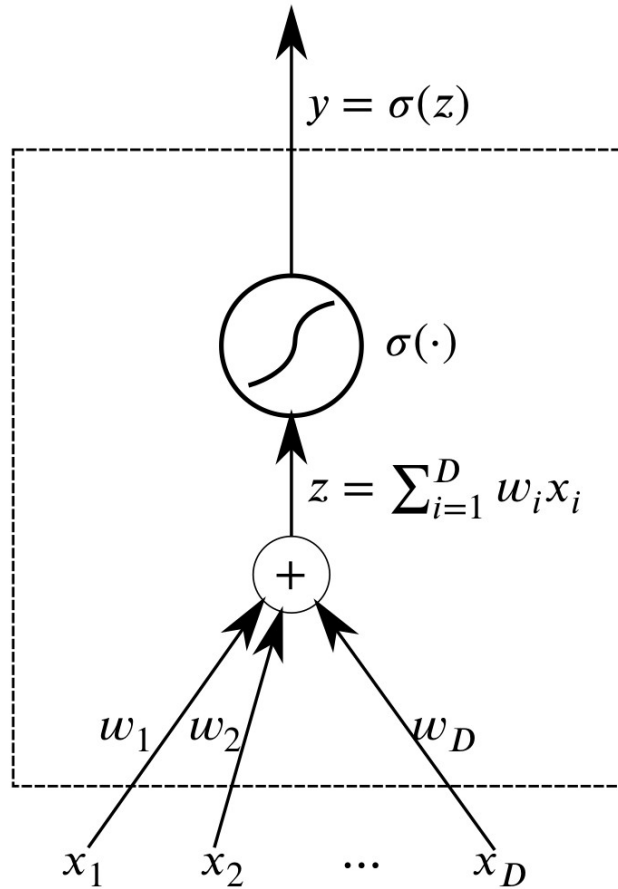
Serialize in SATB order



START	(.)
(62, True)	(60, False)
(56, True)	(57, False)
(52, True)	(52, False)
(47, True)	(50, False)
(60, False)	(.)
(56, True)	(60, True)
(52, True)	(57, True)
(47, True)	(52, True)
	(50, True)

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

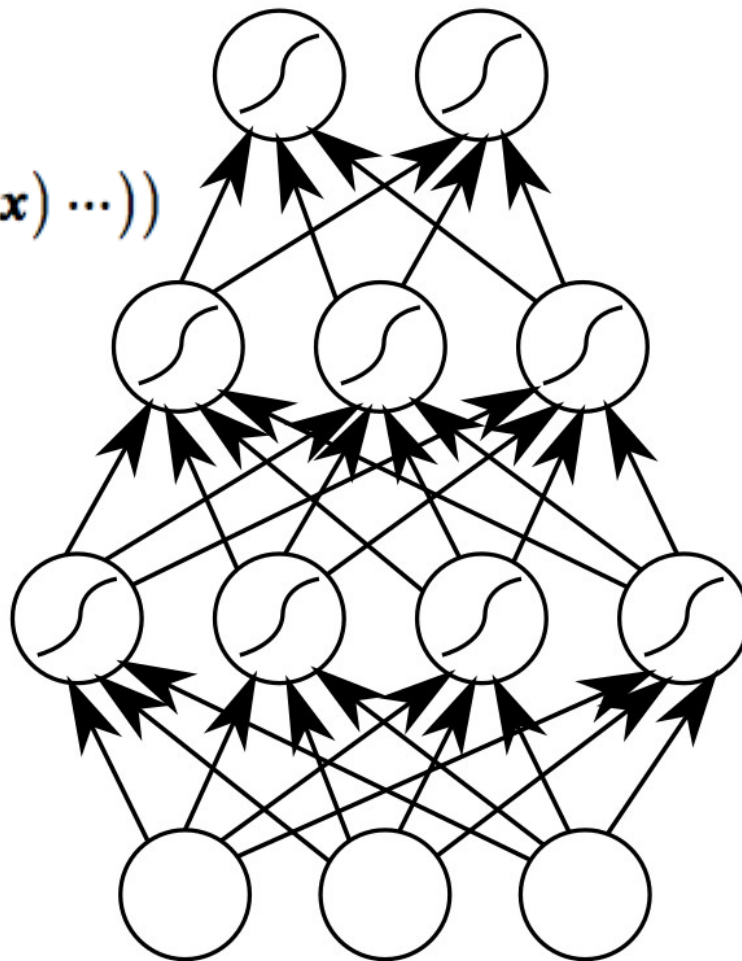
Recurrent neural networks (RNNs)



Neuron

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

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



Output Layer

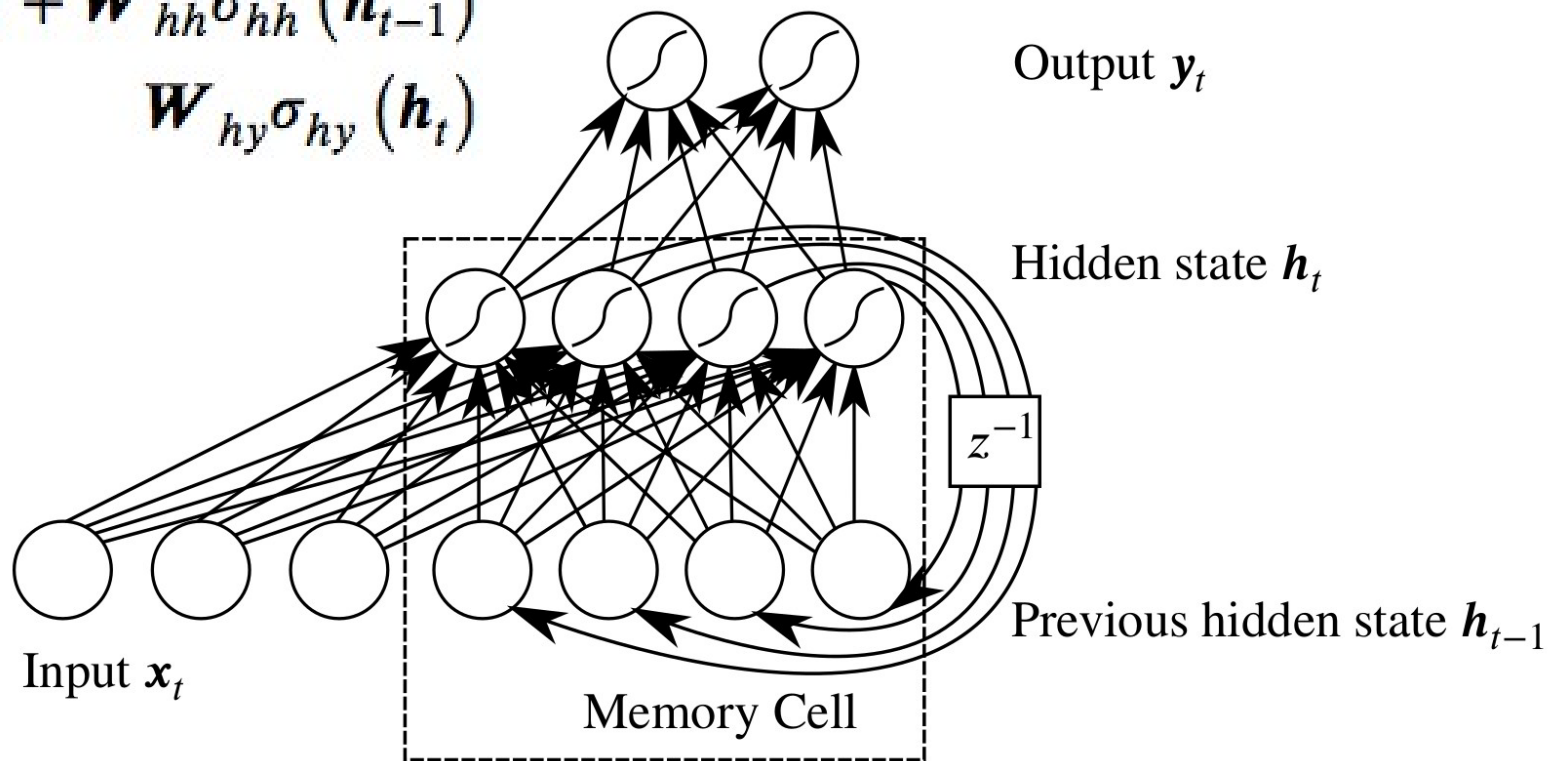
Hidden Layers

Input Layer

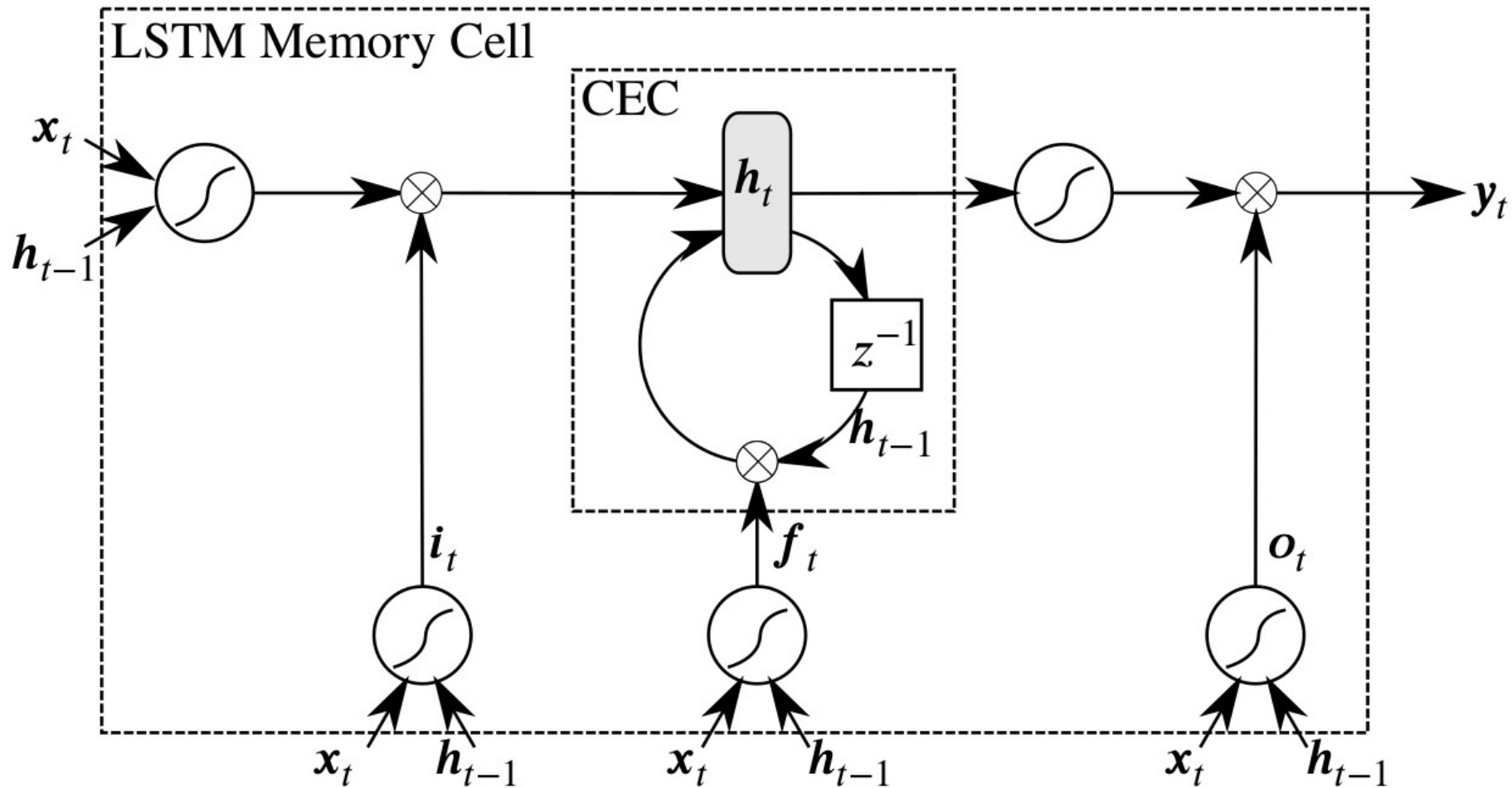
Feedforward neural network

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

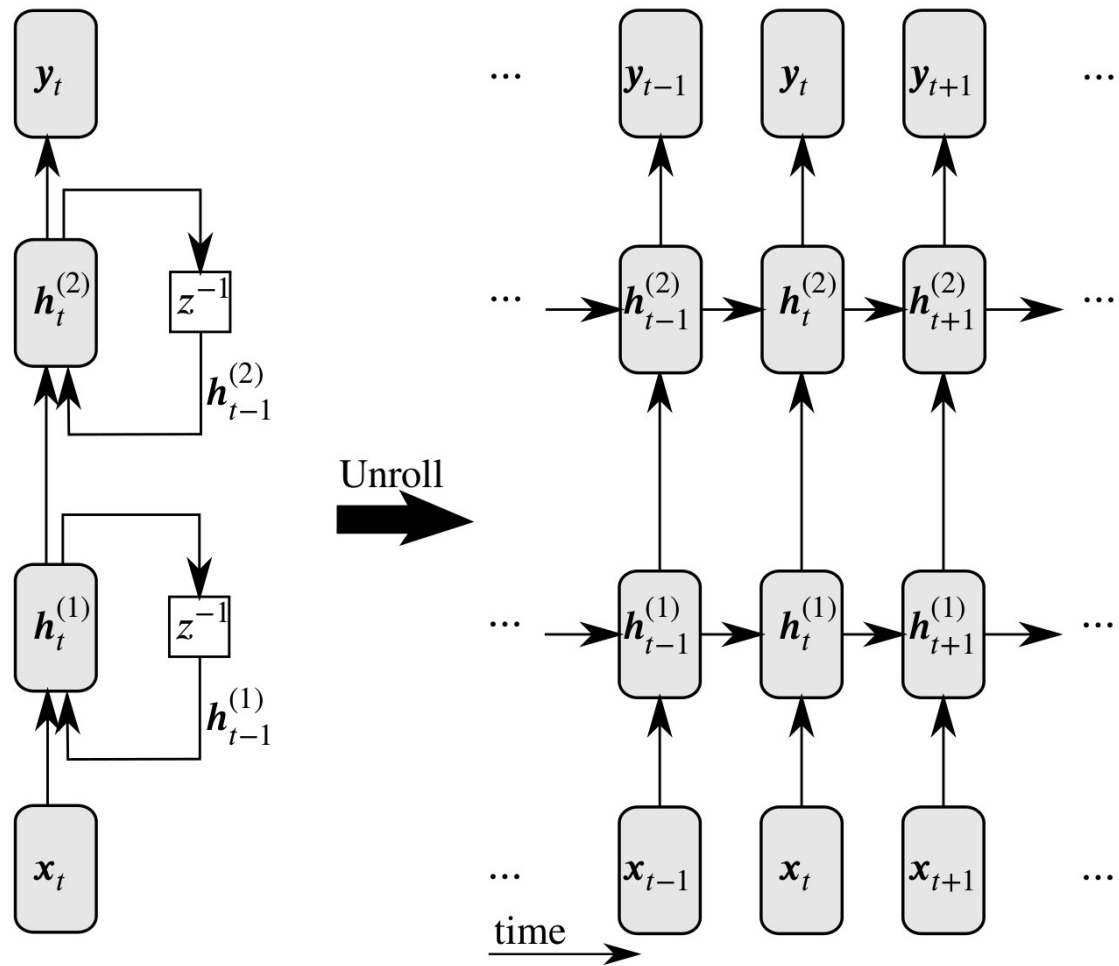
$$\mathbf{y}_t = \mathbf{W}_{hy} \sigma_{hy}(\mathbf{h}_t)$$



Memory cell

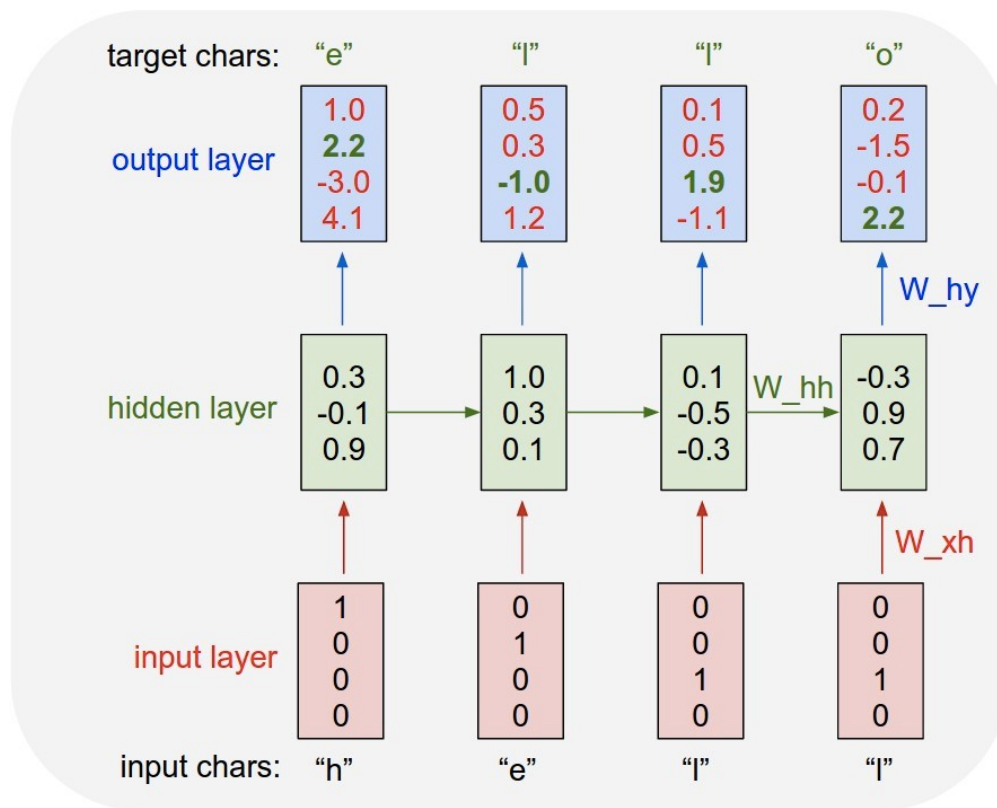


Long short-term memory (**LSTM**) cell



Stacking memory cells to form a deep RNN
Unrolling for training

Sequential prediction training criteria



Training and optimizing BachBot

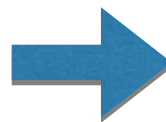
Training BachBot

RNN dynamics

$$\tilde{P}(\mathbf{x}_{t+1} | \mathbf{x}_t, \mathbf{h}_{t-1})$$

Initial state (all 0s)

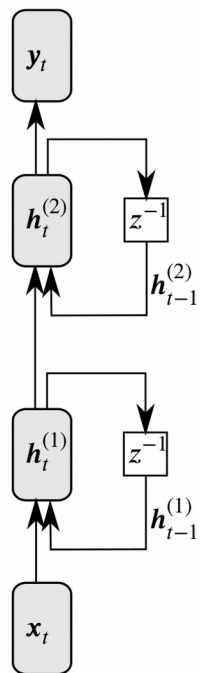
$$+ \mathbf{h}_0$$



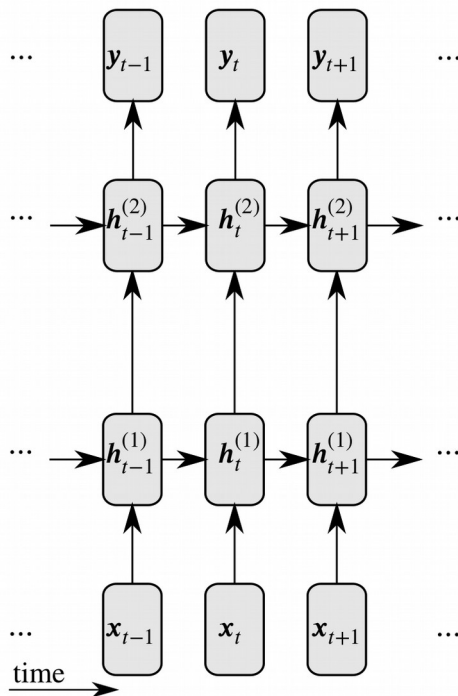
Prob. distr. over sequences

$$\mathbf{y}_t = \tilde{P}(\mathbf{x}_t | \mathbf{x}_{1:t-1})$$

Optimize the
RNN parameters...



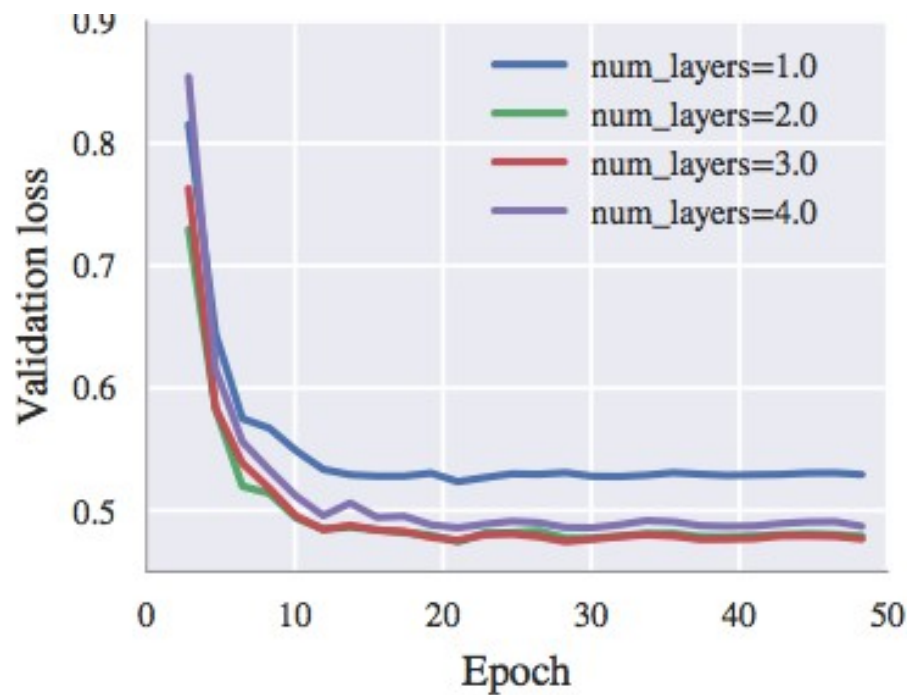
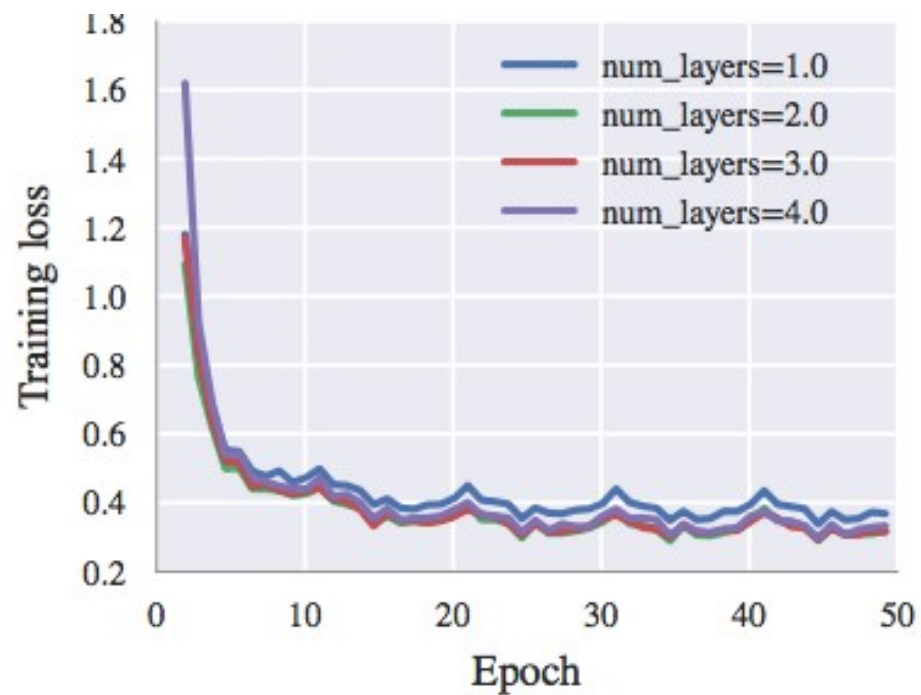
Unroll



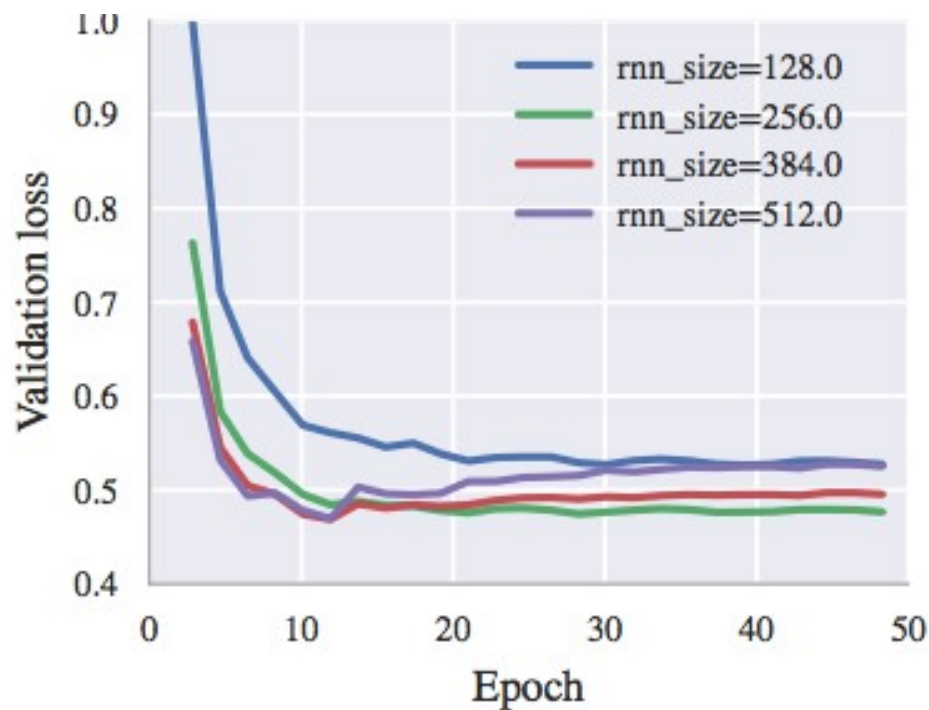
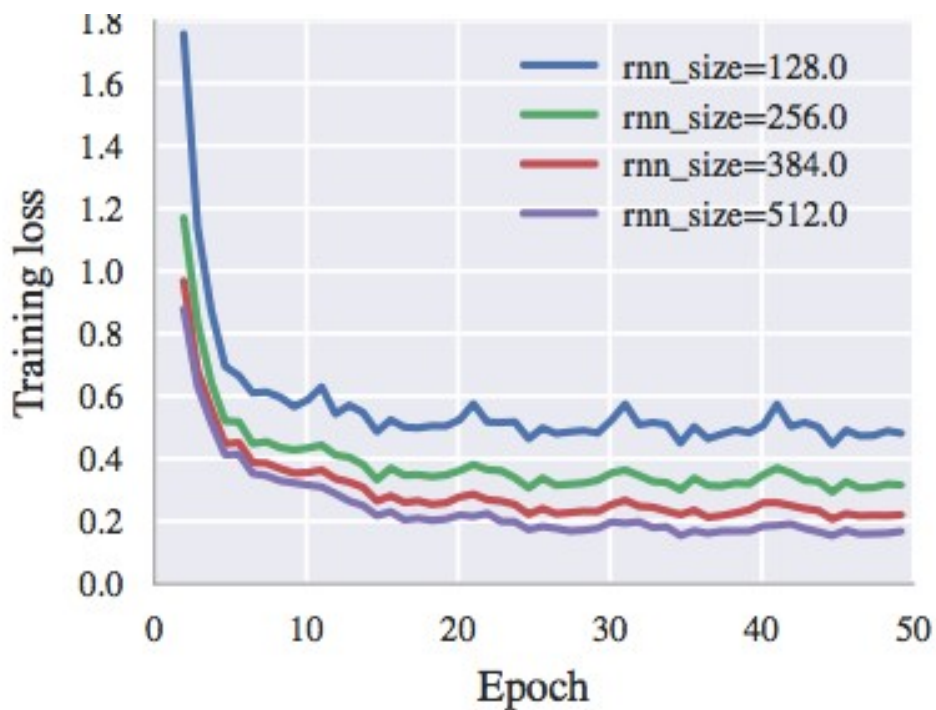
...in order to maximize the
probability of the real
Bach chorales.

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

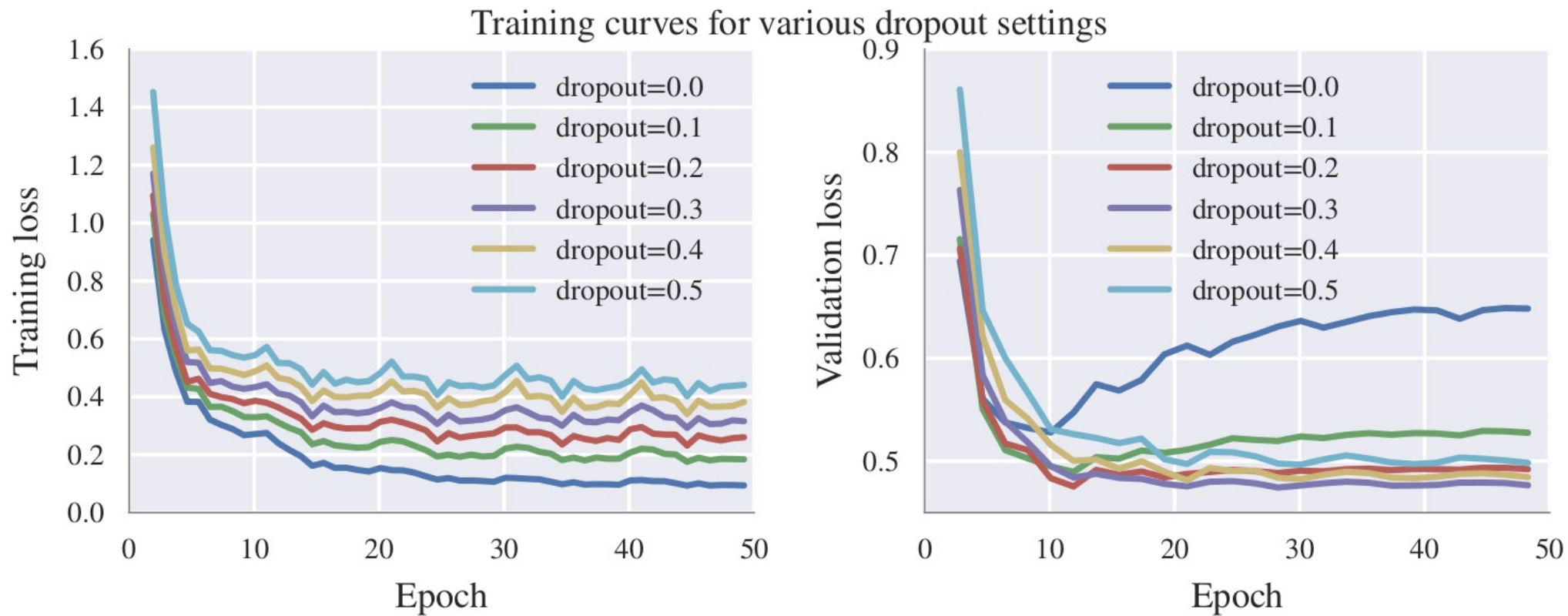
GPUs deliver a 8x performance speedup



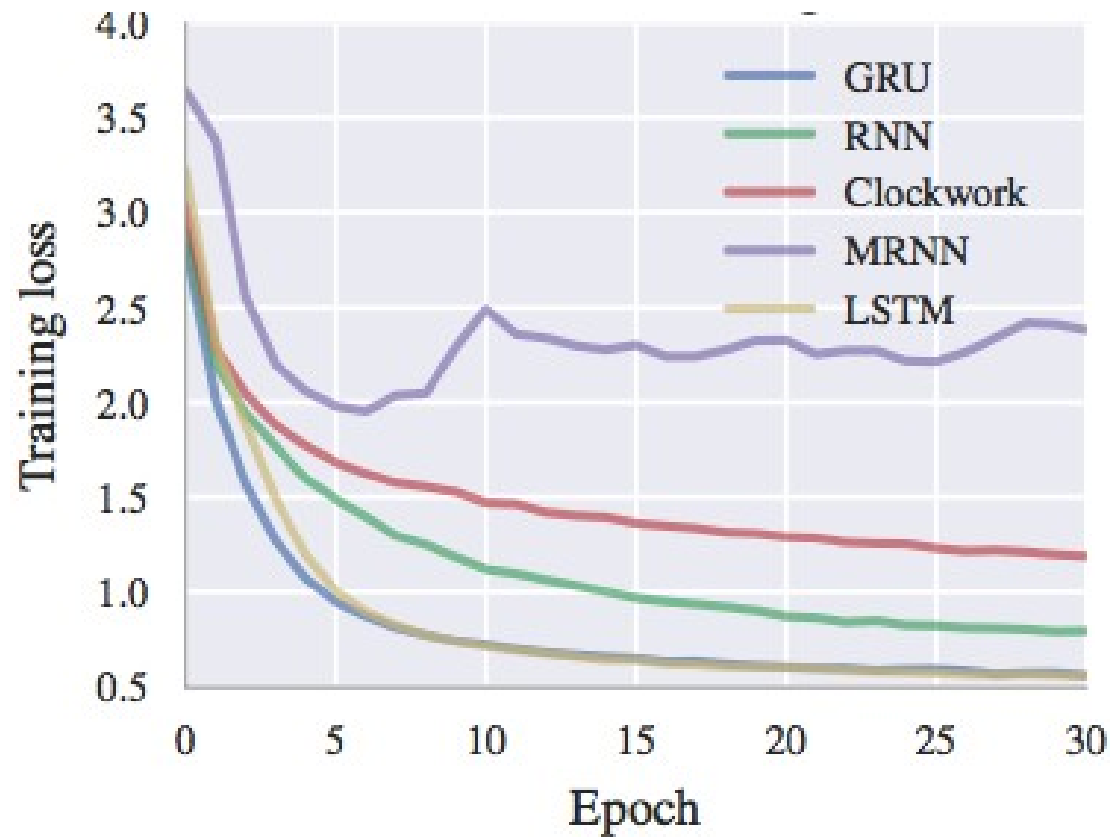
Depth matters (to a certain point)



Hidden state size matters (to a certain point)



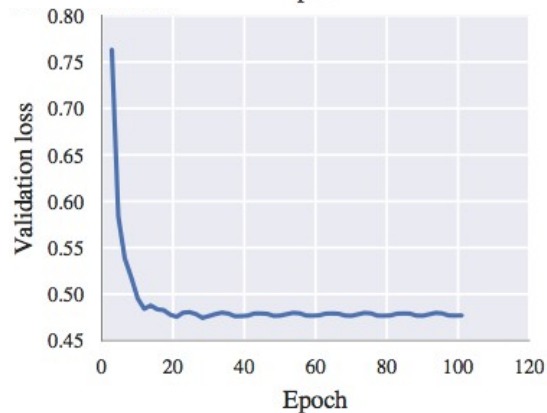
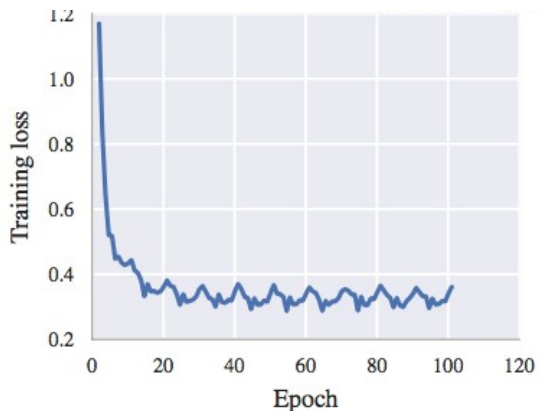
Dropout improves generalization (to a certain point)



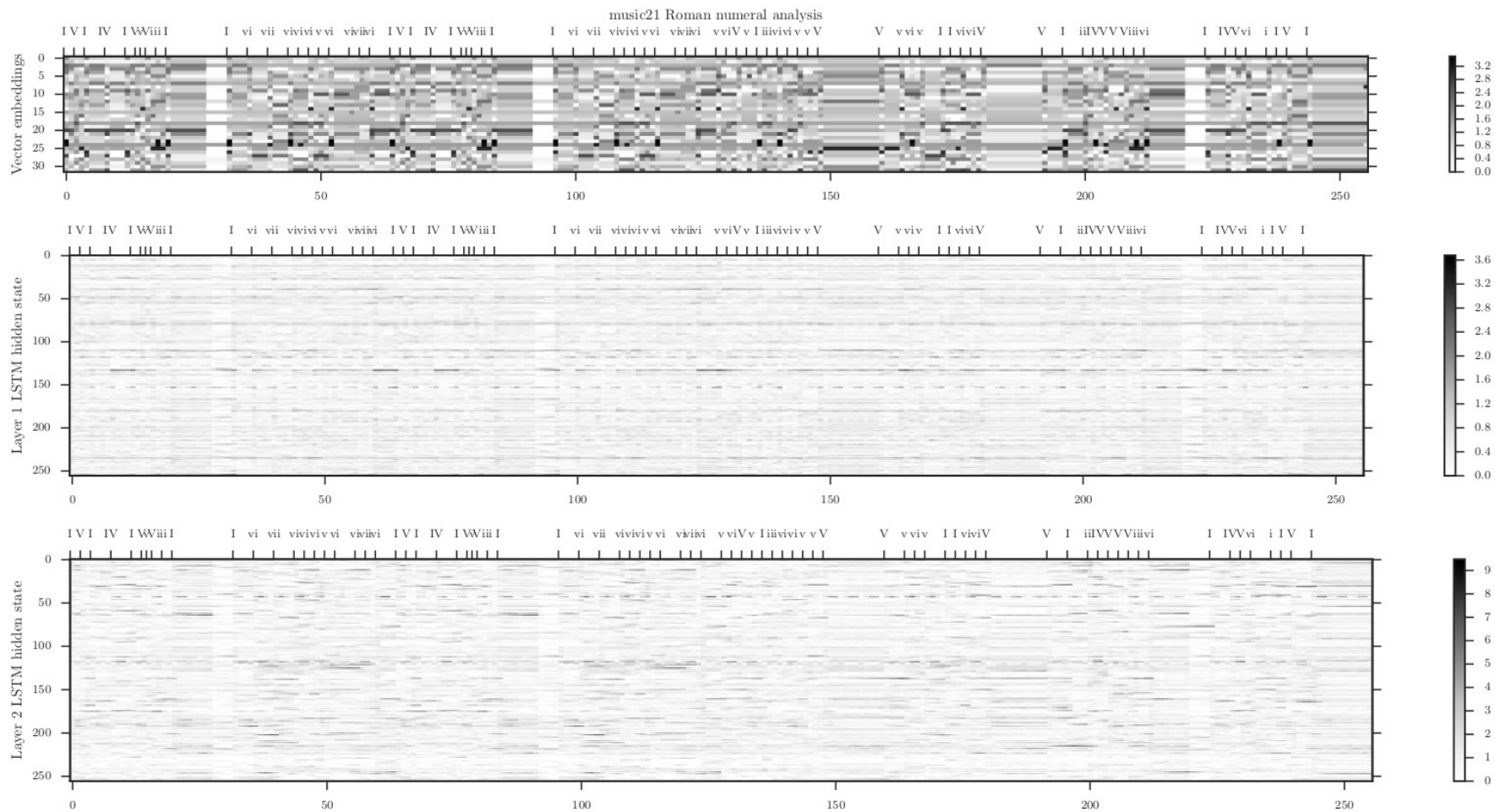
LSTM memory cells matter

The final model

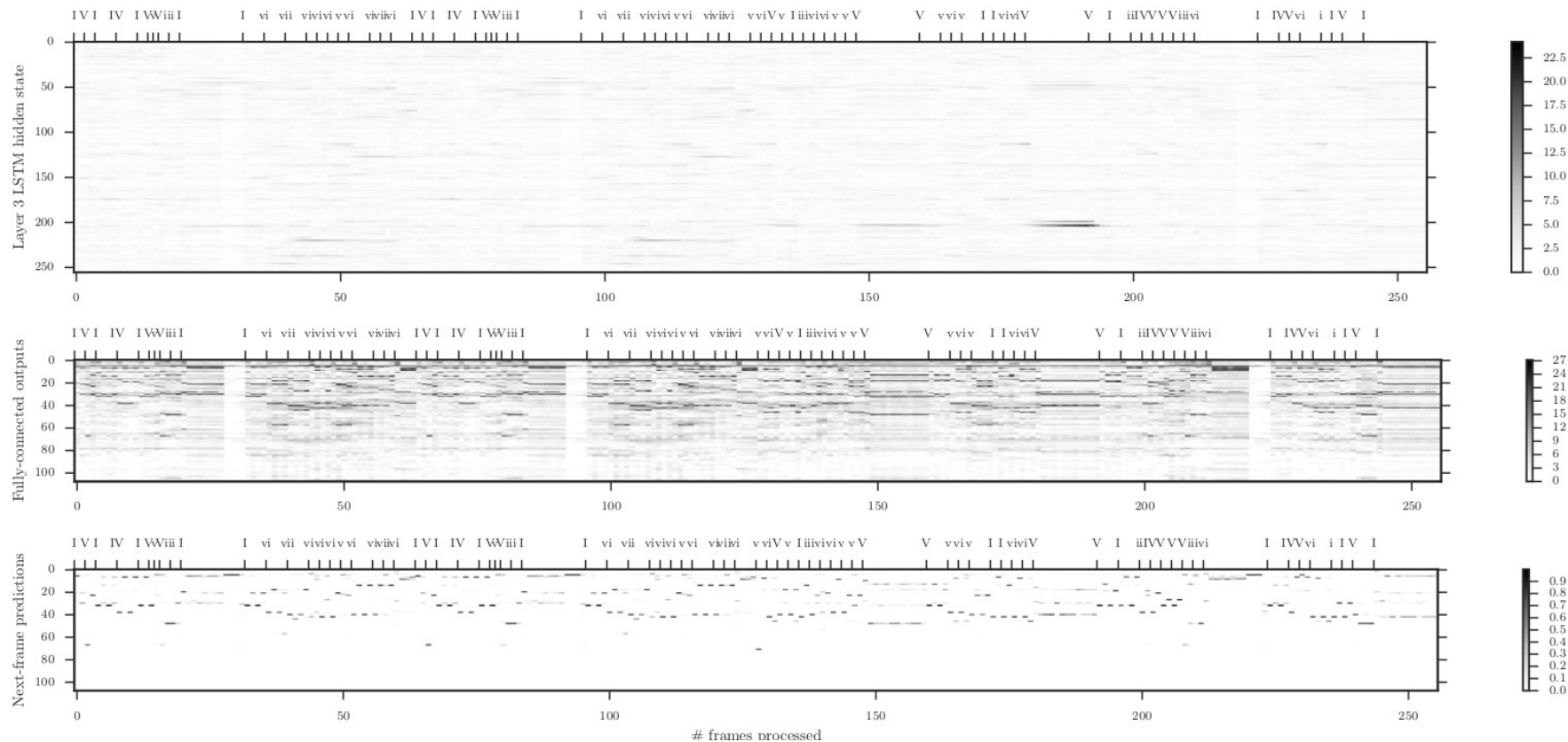
- Architecture
 - 32-dimensional vector space embedding
 - 3 layered, 256 hidden unit LSTMs
- Standard deep learning tricks
 - 30% dropout
 - Batch normalization
 - 128 timestep truncated back-propagation through time (BPTT)
- Bach chorales specific tricks
 - SATB order: S determines harmonic context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



Results



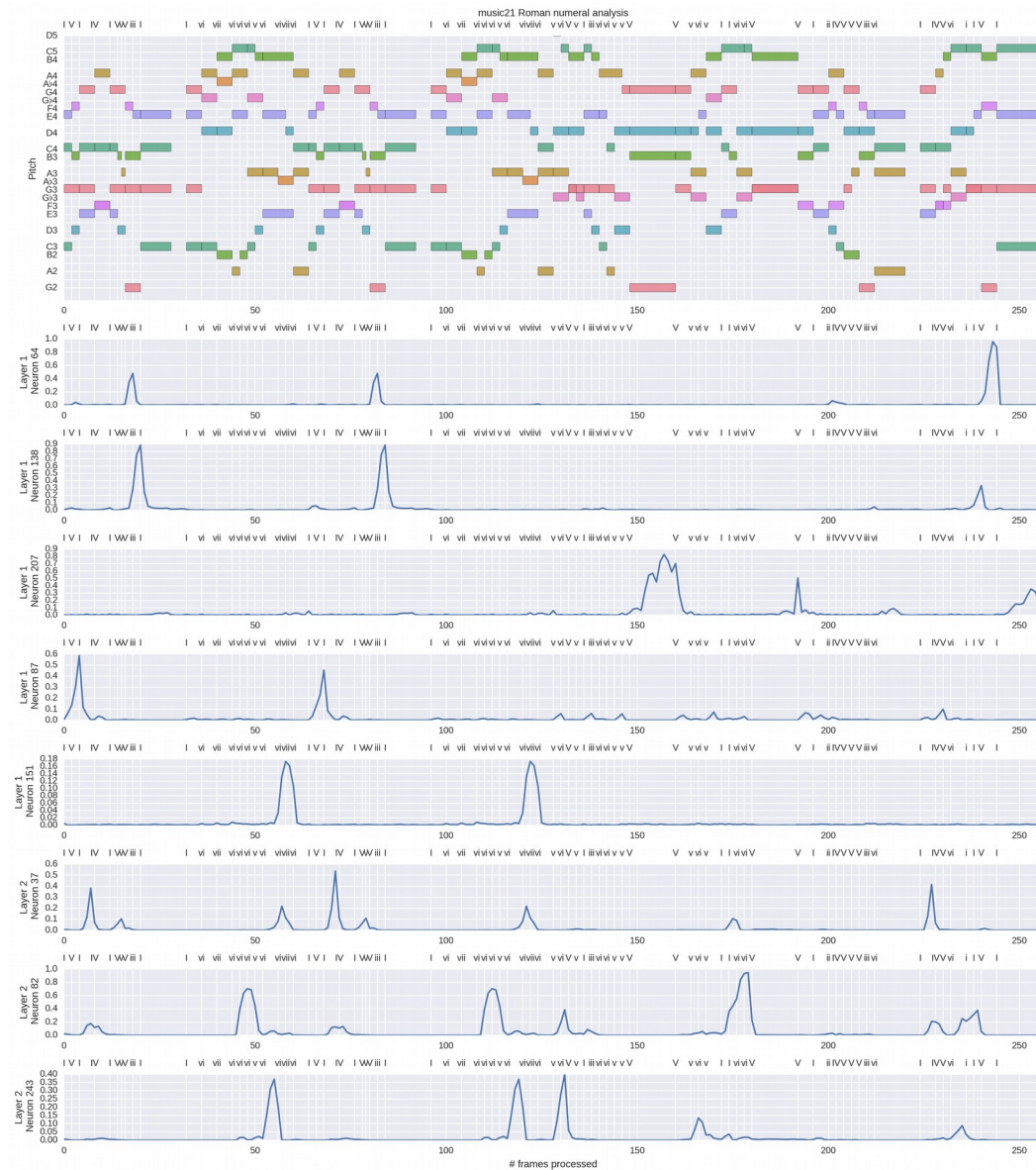
Hidden state is difficult to interpret!
Input and memory cell (layer 1 and 2)

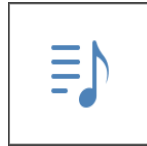


Layers closer to the output resemble piano roll (consequence of sequential training criteria)
Memory cell (layer 3), outputs, and predictions

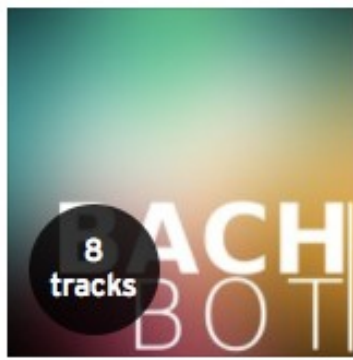
Model learns music theory!

- 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)



BachBot
BachBot.com




	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
	6 Sample-4	▶ 558
	7 Sample-5	▶ 538
	8 C major scale	▶ 828
View fewer tracks		

Like Share More

Harmonization: produce the accompanying parts
<https://soundcloud.com/bachbot>



	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>

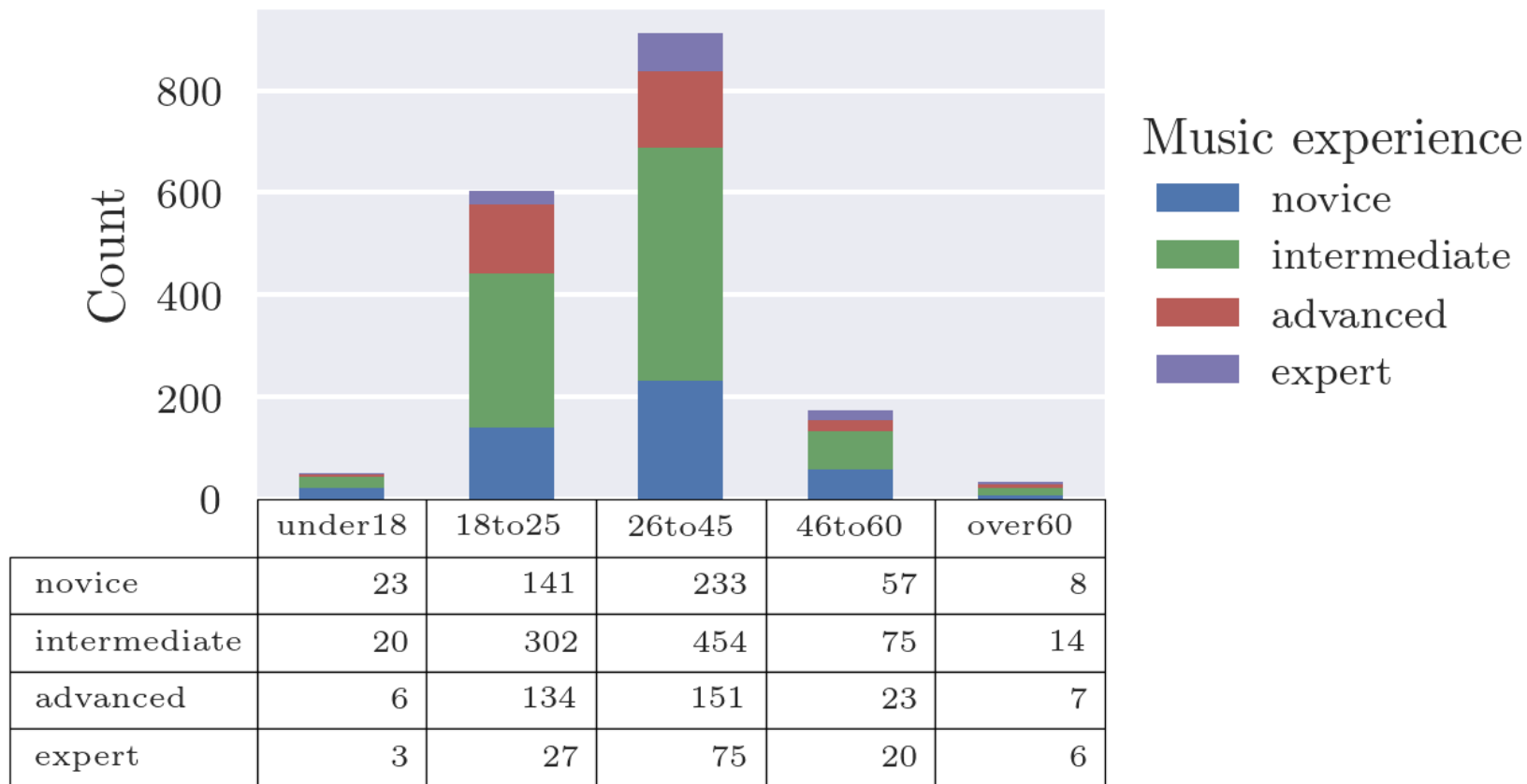
The BachBot Challenge

	Select the music most similar to Bach
Select	  
Select	 
	<div>Submit</div>

20%

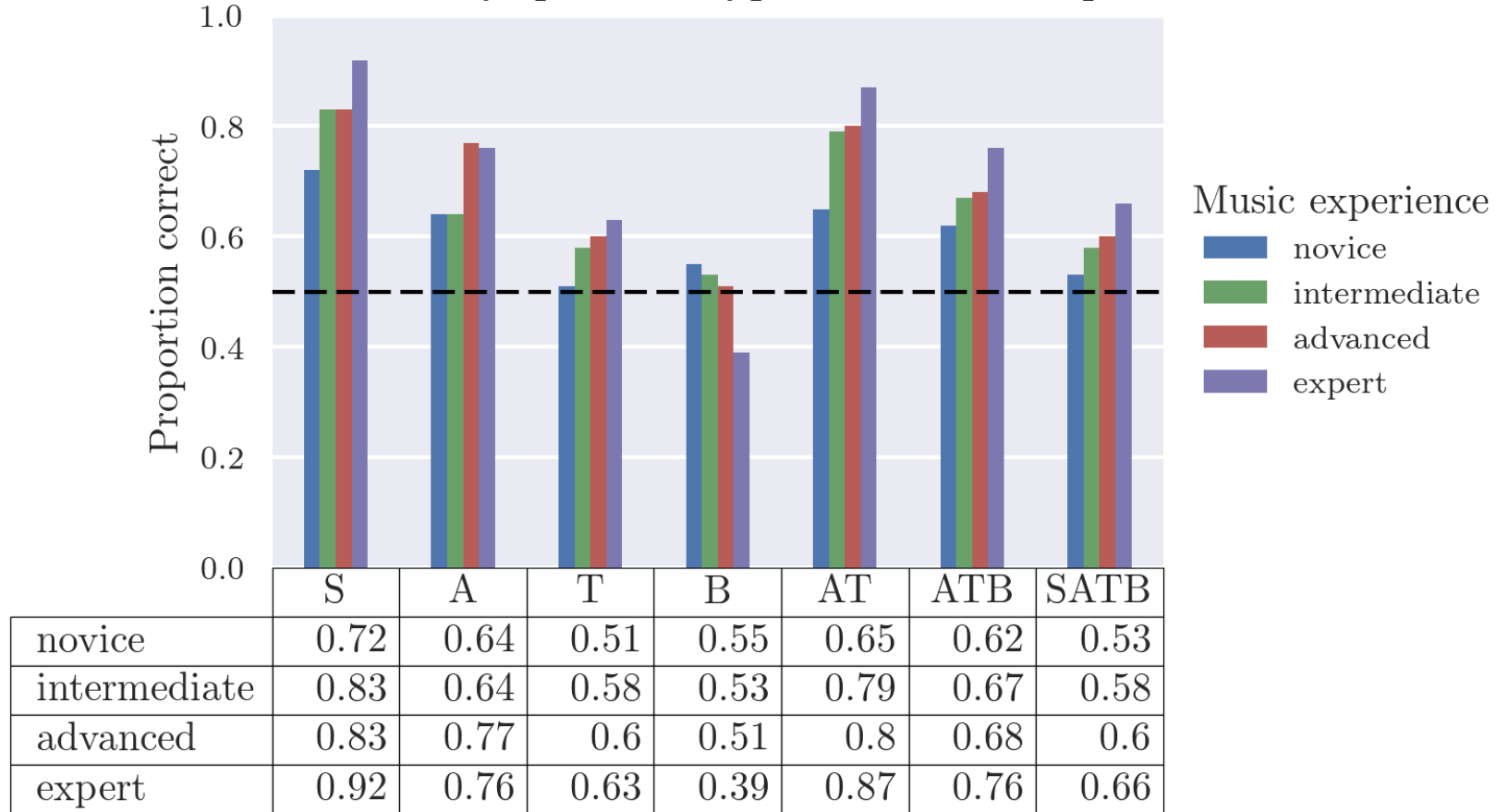
Question 1 out of 5

Participant demographics



Participants by age group and music experience

Performance by question type and music experience



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!

