

Chord Recognition :

홍석준, 조승연, 김윤기, 김성운,
정서현

경소현, 한민석, 심성호, 강성식

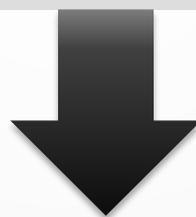


Our goal

**Given
Audio DATA :**



Gm7

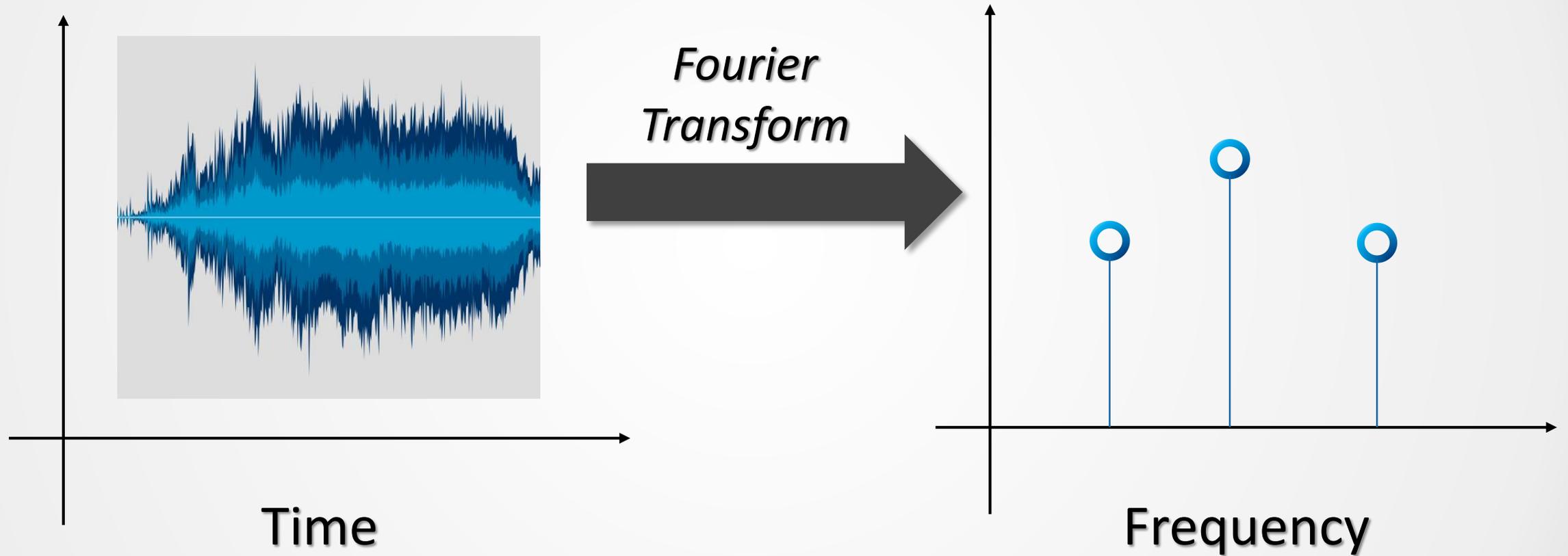


C7

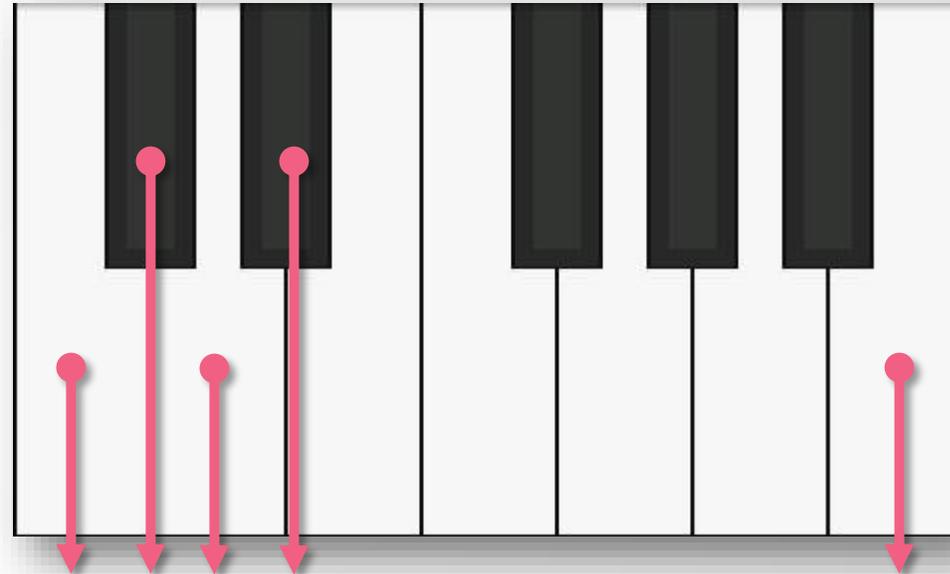


FM7

Feature Extraction



Feature Extraction



Chroma Vector : [1 , 2, 4, 5, ... 10.1]

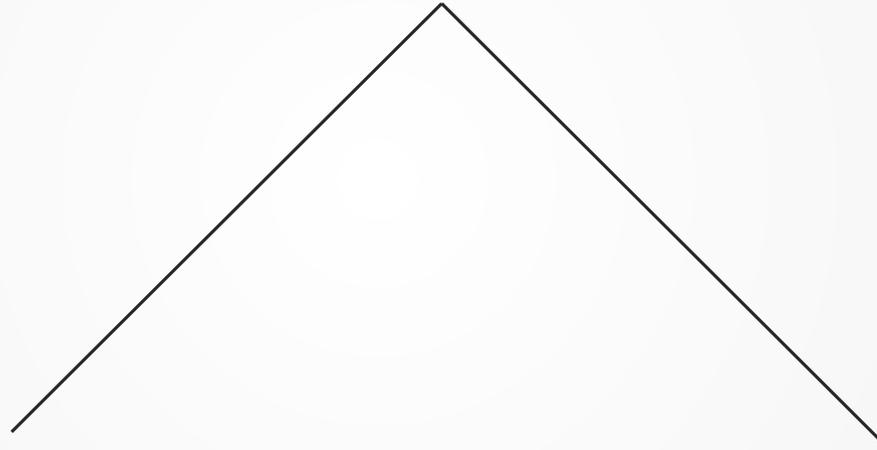
Fujishima (ICMC, 1999)

Mathematical Model

♪ Probabilistic Graphical Model

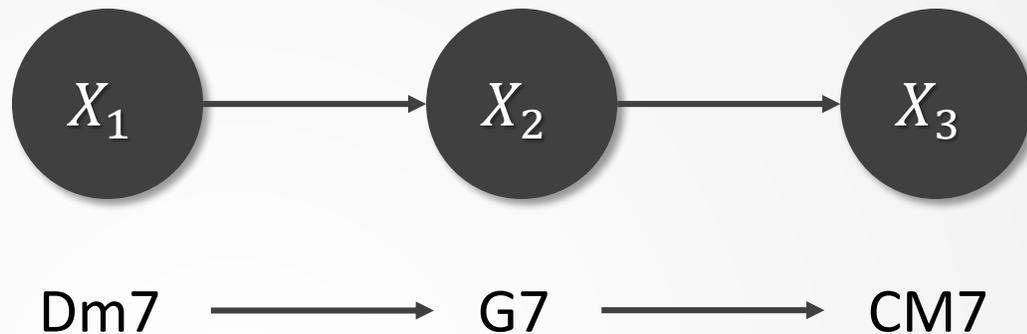
♪ Hidden Markov Model

♪ Bayesian Network

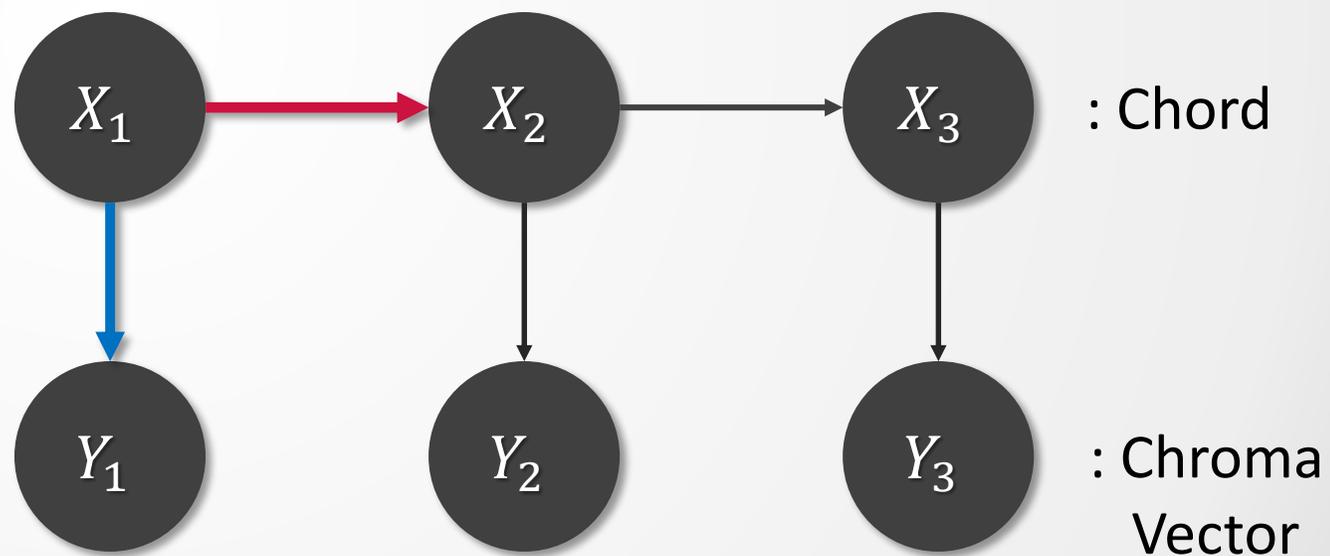


Hidden Markov Model – Sheh, Ellis (2003)

♪ Markov Process



♪ Hidden Markov Model

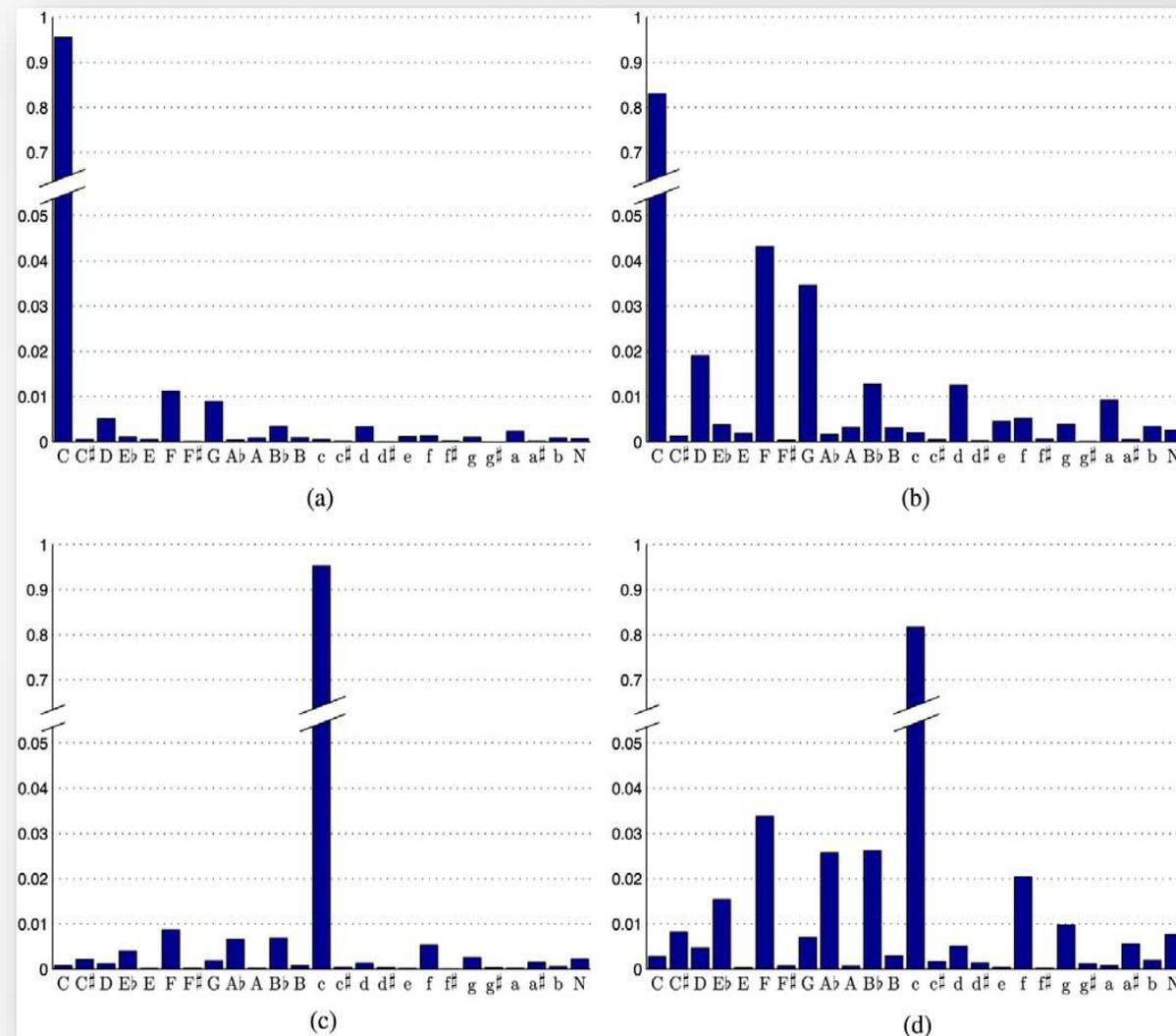


Problems

🎵 Transition Probability

🎵 Emission Probability

🎵 Markov Process

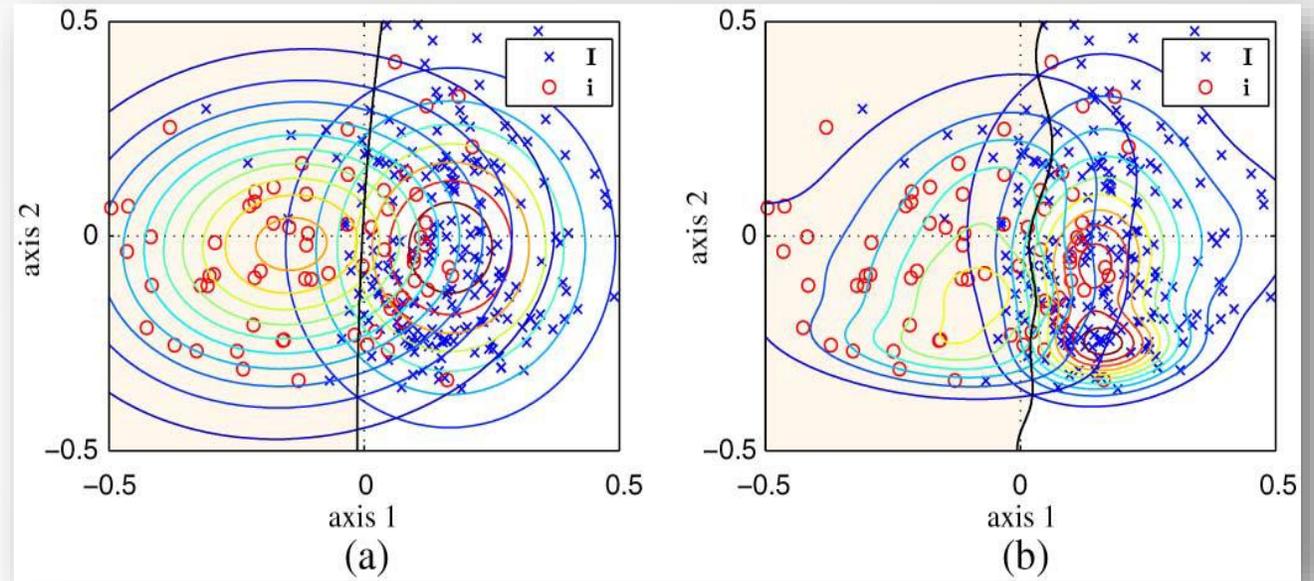


Problems

♪ Transition Matrix

♪ Emission Probability

♪ Markov Process



Problems

♪ Transition Matrix

♪ Emission Probability

♪ Markov Process

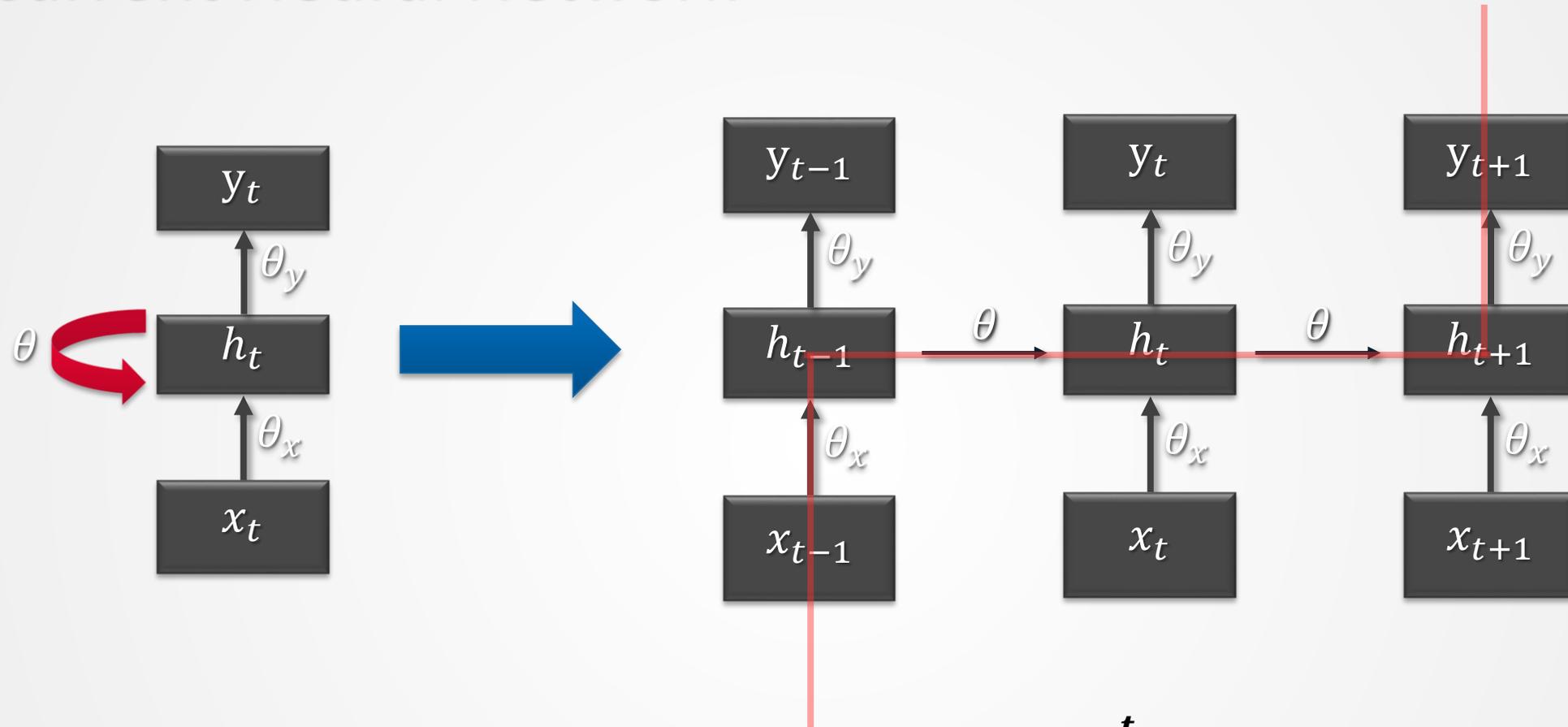
이 등 병 의 편 지

김현성 작사
김현성 작곡
김광석 노래

www.musicscore.co.kr

The image shows a musical score for the song '이 등 병 의 편 지' (Letter from a Soldier). It consists of three staves of music in treble clef with a key signature of one flat (B-flat major). The lyrics are written below the notes. Chord progressions are indicated above the notes. The first staff shows the beginning of the melody with chords C, E, Am, Am/G, Dm, and G. The second staff contains the main body of the song with lyrics: '나 와 - 열 - 차 타 고 혼 련 소 로 가 는 날 부 모 들 아 - 군 - 대 가 면 편 지 꼭 해 다 오 그 대 잘 린 - 내 - 머 리 가 처 음 에 는 우 습 다 가 거 을'. The third staff continues the melody with lyrics: '님 께 큰 절 하 고 대 문 밖 을 나 실 때 가 습 들 과 과 즐 거 왔 던 문 밖 을 나 실 때 가 습 속 에 비 친 모 습 이 군 어 진 다 마 음 까 지 지 슴 차 동'. Chord progressions for the second and third staves include C, E, Am, Dm, G7, and C.

Recurrent Neural Network



$$w = w - \alpha \frac{\partial E}{\partial w} \quad (1)$$

$$\frac{\partial h_t}{\partial h_k} = \prod_{i=k+1}^t \frac{\partial h_i}{\partial h_{i-1}} \quad (2)$$

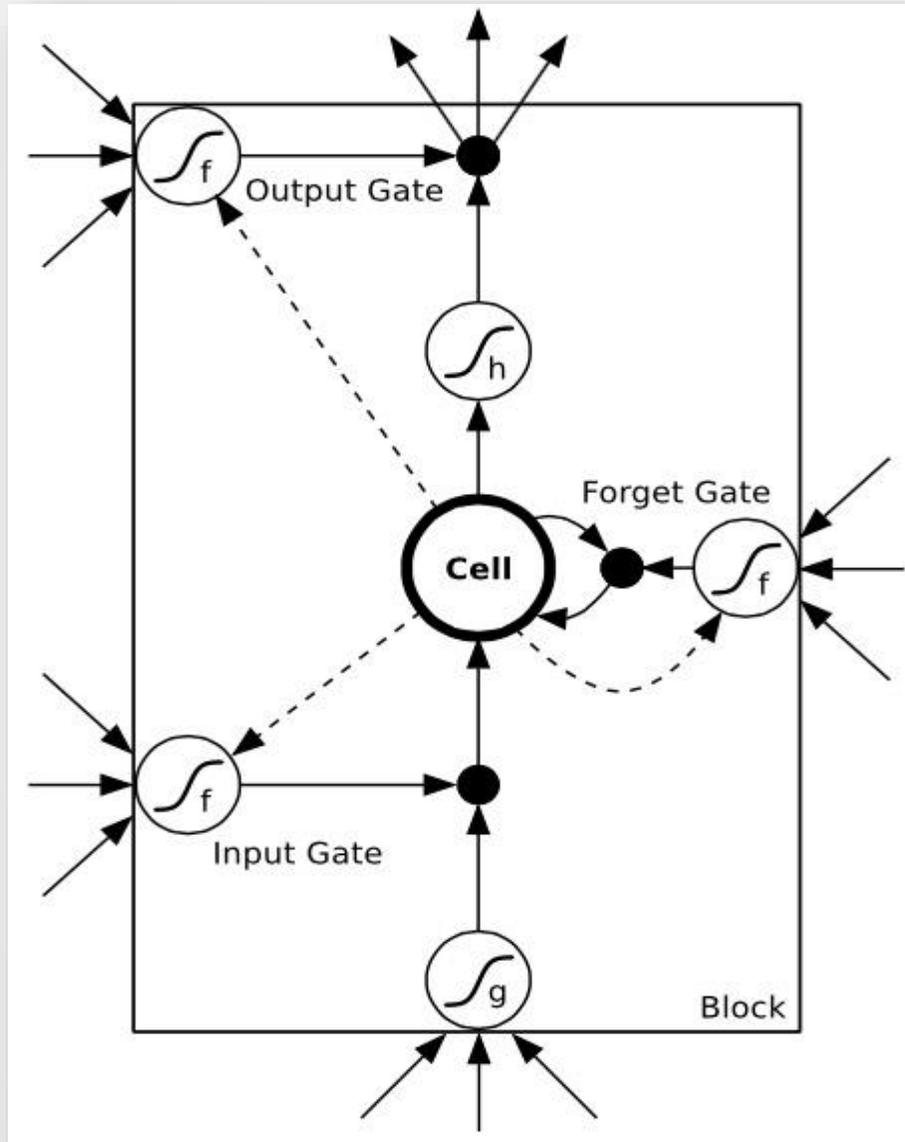
Vanishing Gradient Problem

$$\frac{\partial h_t}{\partial h_k} = \prod_{i=k+1}^t \frac{\partial h_i}{\partial h_{i-1}}$$

if $\left\| \frac{\partial h_i}{\partial h_{i-1}} \right\| < \mathbf{1}$, then $\left\| \frac{\partial h_t}{\partial h_k} \right\| \rightarrow 0$ for some cases.

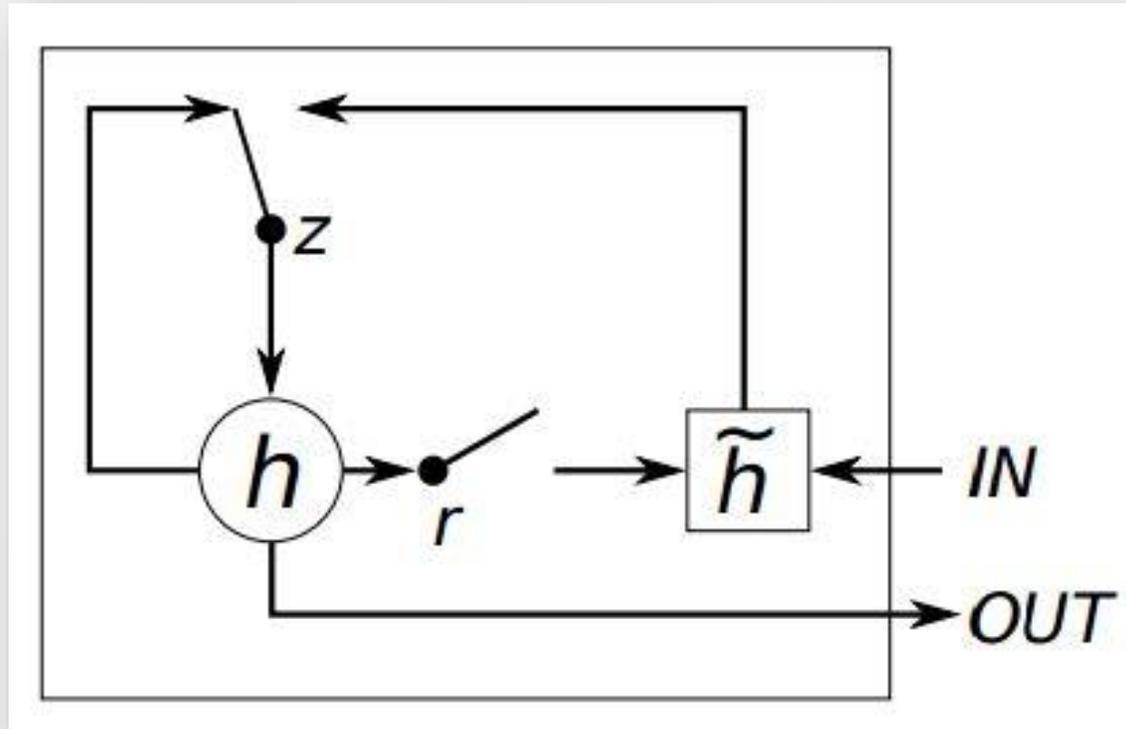
if $\left\| \frac{\partial h_i}{\partial h_{i-1}} \right\| > \mathbf{1}$, then $\left\| \frac{\partial h_t}{\partial h_k} \right\| \rightarrow \infty$ for some cases.

LSTM (Long Short-term Memory)



Schmidhuber, Hochreiter
"LONG SHORT-TERM
MEMORY" (1997)

GRU (Gated Recurrent Neural Networks)



Chung, Junyoung, et al.
“Empirical evaluation of
gated recurrent neural
networks on sequence
modeling.” (2014)

Data set - Beatles

이름	수정한 날짜
 A_Hard_Day_s_Night	2016-01-31 오후 2...
 Abbey_Road	2016-01-31 오후 2...
 Beatles_For_Sale	2016-01-31 오후 2...
 Help_	2016-01-31 오후 2...
 Let_It_Be	2016-01-31 오후 2...
 Magical_Mystery_Tour	2016-01-31 오후 2...
 Please_Please_Me	2016-01-31 오후 2...
 Revolver	2016-01-31 오후 2...
 Rubber_Soul	2016-01-31 오후 2...
 Sgt_Pepper_s_Lonely_Hearts_Club_Band	2016-01-31 오후 2...
 The_White_Album_Disc_1	2016-01-31 오후 2...
 The_White_Album_Disc_2	2016-01-31 오후 2...
 With_The_Beatles	2016-01-31 오후 2...
 MODELS.mat	2010-04-07 오전 7...

MIR Research

TUESDAY, 25 SEPTEMBER 2007

Beatles Chord Transcriptions

Chris Harte from the [C4DM](#) announced last night that he completed his amazing effort of transcribing the chords for all songs on the 12 studio albums of the Beatles. The transcriptions are extremely high quality. Anyone who wants a copy just needs to contact him. This will definitely boost research in any direction related to chords (chord recognition, chord progressions, harmony analysis...). It's also a good excuse for any research lab to buy the complete Beatles collection. Btw, don't forget to cite his [work](#) when you use his annotations! :-)



Below are excerpts from the two emails he sent to the music-ir mailing list, so that Google can index them (Afaik he hasn't set up a website for this yet).

(Chris' email is christopher dot harte at elec dot qmul dot ac dot uk).

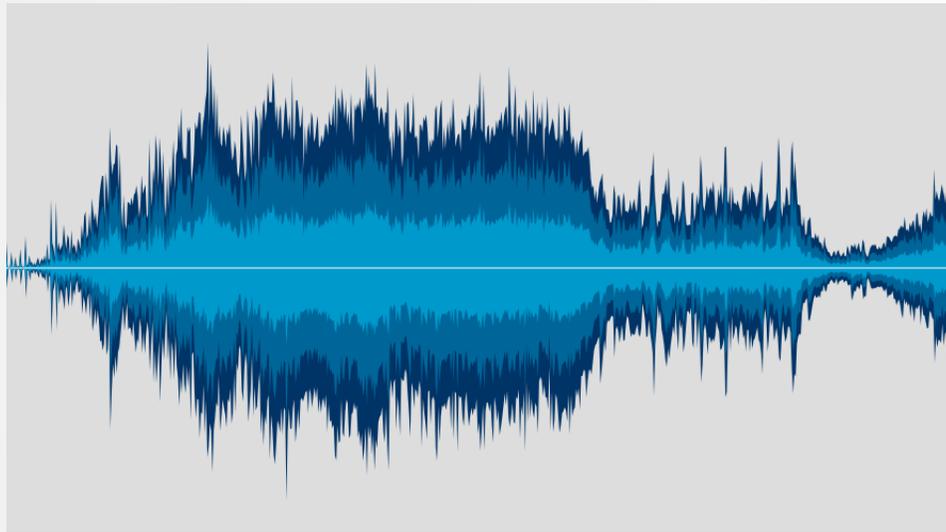
Chris Harte wrote in his first email (Sep 24, 2007, 9pm):

[...] I have just completed work on the full set of chord transcriptions for the beatles songs from all 12 studio albums.

ARCHIVE

- ▶ 2010 (2)
- ▶ 2009 (11)
- ▶ 2008 (43)
- ▼ 2007 (61)
 - ▶ [December](#) (1)
 - ▶ [October](#) (1)
 - ▼ [September](#) (15)
 - [ISMIR Highlights](#)
 - [Beatles Chord Transcriptions](#)
 - [One Llama, IMIRSEL, MIREX](#)
 - [ISMIR Highlight: Recommendation Tutorial](#)
 - [Fun things to do with fingerprinting](#)
 - [The most frequently cited ISMIR paper](#)
 - [One evening and no testing](#)
 - [Marsyas & Music Classification](#)
 - [Overfitting and MIREX](#)
 - [Vocaloid 2 is a big hit in Japan](#)

Goal



C

C#

D

D#

E

⋮

Cm

C#m

Dm

D#m

Em

N

HMM case...

57.7%

Results - RNN

Epoch 36/40

10s - loss: 0.5620 - acc: 0.8327 - val_loss: 0.7360 - val_acc: 0.7900

Epoch 37/40

10s - loss: 0.5627 - acc: 0.8328 - val_loss: 0.7552 - val_acc: 0.7803

Epoch 38/40

10s - loss: 0.5603 - acc: 0.8331 - val_loss: 0.7687 - val_acc: 0.7783

Epoch 39/40

10s - loss: 0.5622 - acc: 0.8324 - val_loss: 0.8207 - val_acc: 0.7871

Epoch 40/40

10s - loss: 0.5676 - acc: **0.8328** - val_loss: 0.7631 - val_acc: 0.7865

Train on **73000** samples, validate on **3072** samples

test_SimpleRNN

[0.66649848595261574, 0.80436197916666663]

Results - LSTM

Epoch 36/40

27s - loss: 0.5398 - acc: 0.8553 - val_loss: 0.7132 - val_acc: 0.8213

Epoch 37/40

27s - loss: 0.5469 - acc: 0.8544 - val_loss: 0.8049 - val_acc: 0.7910

Epoch 38/40

27s - loss: 0.5557 - acc: 0.8559 - val_loss: 0.7751 - val_acc: 0.7881

Epoch 39/40

27s - loss: 0.5481 - acc: 0.8552 - val_loss: 0.7311 - val_acc: 0.8070

Epoch 40/40

27s - loss: 0.5567 - acc: **0.8527** - val_loss: 0.7873 - val_acc: 0.8024

test_LSTM

[0.72307247652982676,0.81510416666666663]]

Results - GRU

Epoch 36/40

25s - loss: 0.3815 - acc: 0.8887 - val_loss: 0.5499 - val_acc: 0.8486

Epoch 37/40

25s - loss: 0.3882 - acc: 0.8898 - val_loss: 0.5488 - val_acc: 0.8597

Epoch 38/40

25s - loss: 0.3855 - acc: 0.8898 - val_loss: 0.5435 - val_acc: 0.8607

Epoch 39/40

25s - loss: 0.3863 - acc: 0.8906 - val_loss: 0.5334 - val_acc: 0.8503

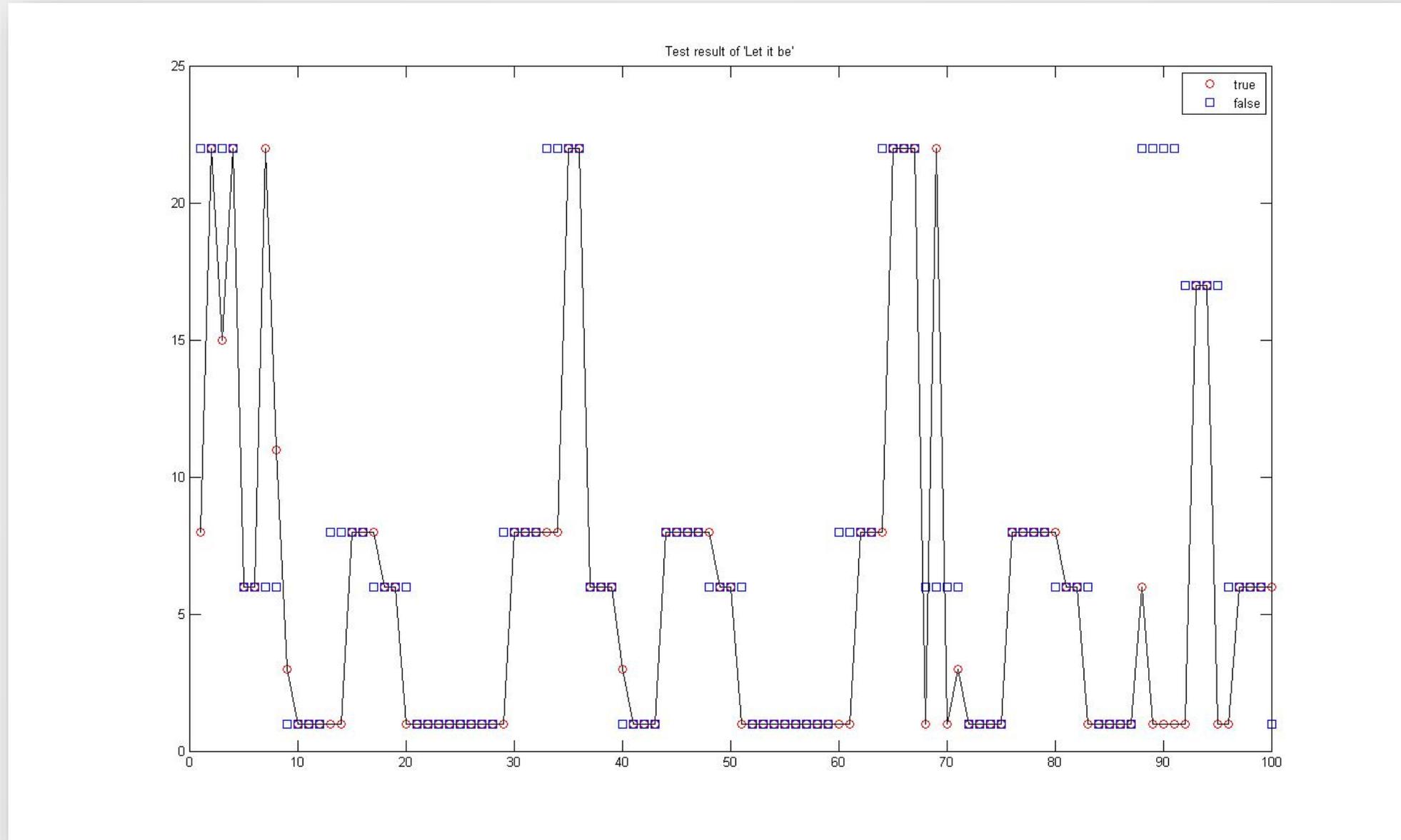
Epoch 40/40

25s - loss: 0.3913 - acc: **0.8894** - val_loss: 0.5160 - val_acc: 0.8590

song_GRU_Let it be

[1.3216885417255002, 0.76119402985074625]

Results – GRU (Let it be)



AUDIO CHORD RECOGNITION WITH RECURRENT NEURAL NETWORKS

Nicolas Boulanger-Lewandowski, Yoshua Bengio and Pascal Vincent

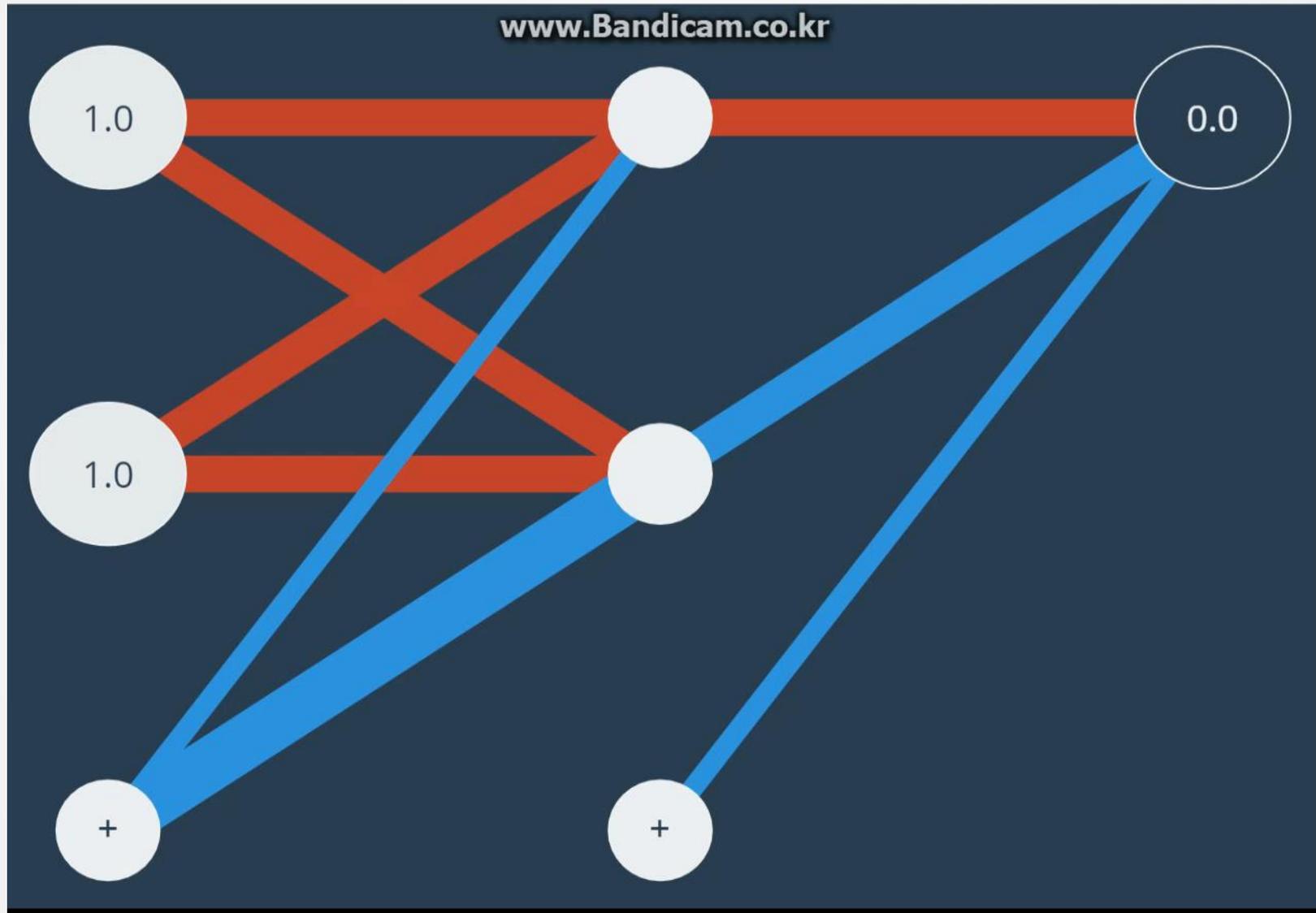
Dept. IRO, Université de Montréal

Montréal, Québec, Canada H3C 3J7

{boulanni, bengioy, vincentp}@iro.umontreal.ca

Method	OR	WAOR
Chordino [22]	80.2%	79.5%
GMM + HMM [20]	82.9%	81.6%
HPA [25]	83.5%	82.7%
Proposed (DBN-2)	89.5%	89.8%
Proposed (RNN)	93.5%	93.6%

Neural Networks



Automatic Composition & Arrangement

CM7

Am7

Dm7

G7

CM7

Am7

Dm7

Lead sheet

Twinkle, Twinkle

The image shows a musical score for the song "Twinkle, Twinkle" in 4/4 time. It consists of three staves of music. The first staff begins with a treble clef, a 4/4 time signature, and a key signature of one flat (B-flat). The melody is written in quarter notes. Above the staff, chord symbols are placed: C, F, C, F, C, G, C. The second staff continues the melody with the same note values. Above it, the chord symbols are: C, F, C, G, C, F, C, G. The third staff concludes the piece with a double bar line. Above it, the chord symbols are: C, F, C, F, C, G, C.

Future work

