

Learning to Play Jazz with Deep Belief Networks

Greg Bickerman, Harvey Mudd College

Sam Bosley, Stanford University

Peter Swire, Brandeis University

Robert Keller, Harvey Mudd College

Motivation

- People are able to improvise jazz on the spot
- Jazz Improvisation
 - Patterned and structured
 - Creative and novel
- Could a machine learn to improvise as well as a human?

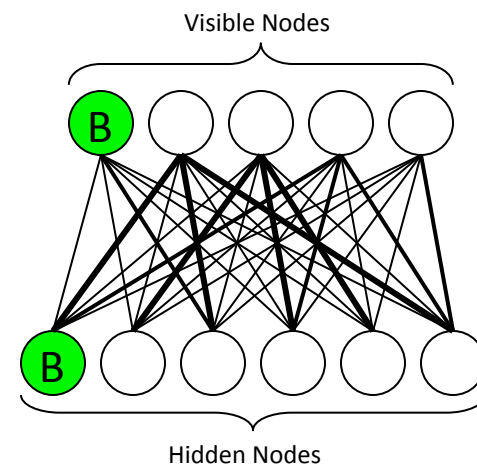


Motivation

- Artificial jazz improvisers already exist
 - GenJam
 - Supervised genetic learning
 - Impro-Visor
 - Extensive musical knowledge built in
- Interested in unsupervised learning
- Minimal representational assumptions

Restricted Boltzmann Machines

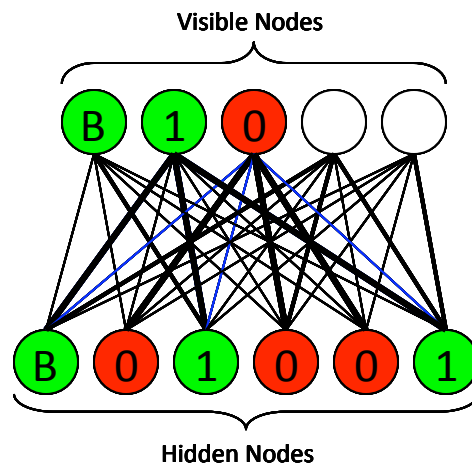
- 2-layer network
 - Visible layer
 - Hidden layer
- Nodes
 - Interconnected
 - Can be set ON or OFF
- Weights
 - Assigned to each connection
 - Symmetric



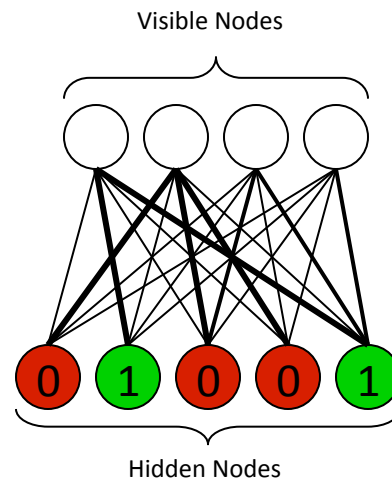
Activation

- Nodes activated **probabilistically** based on activation states of nodes in opposite layer
 - Compute weighted sum of active connections
 - Activation function determines probability of firing

Activation



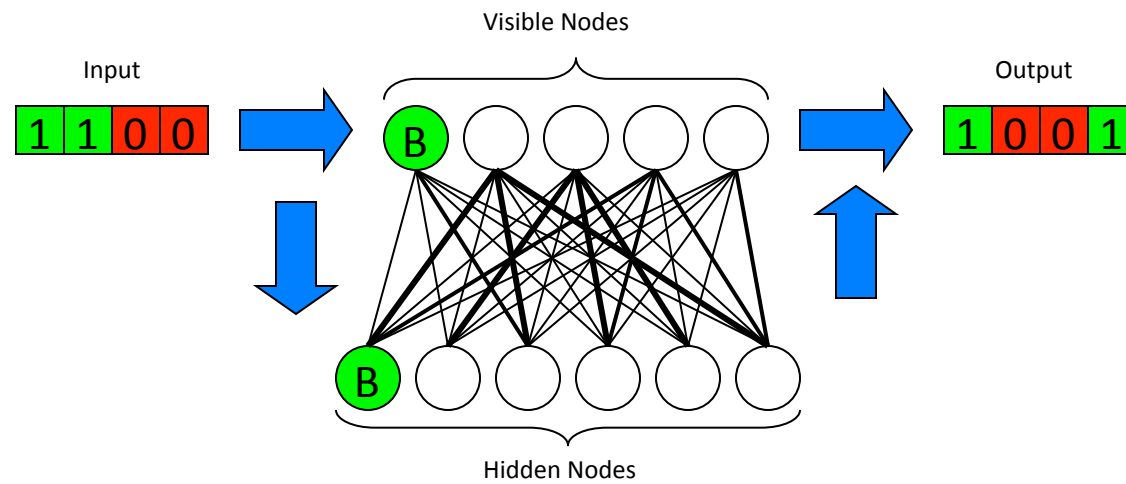
Activation



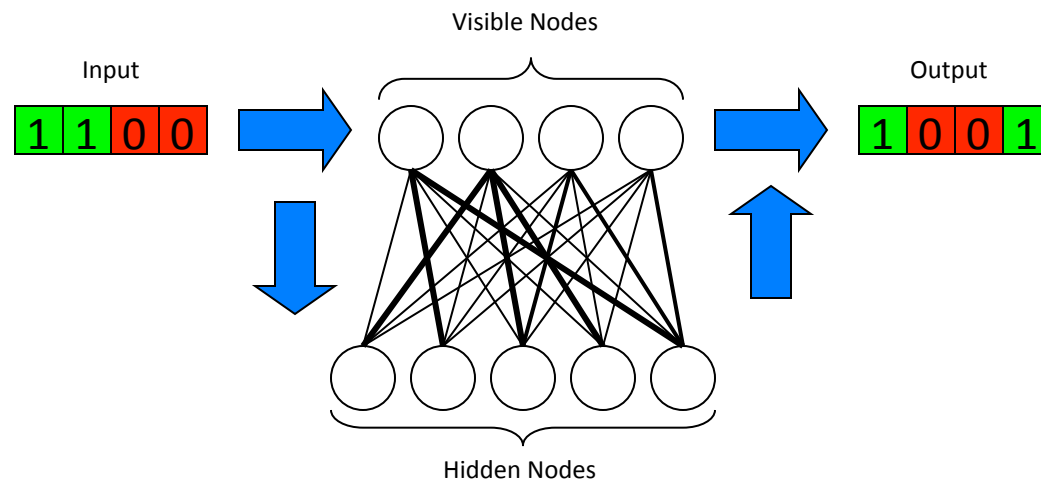
Input / Output

- Input
 - Binary data sequences
 - Mapped onto visible neurons
- Output
 - Identically sized data sequences
 - Read off of visible neurons

Input / Output



Input/Output

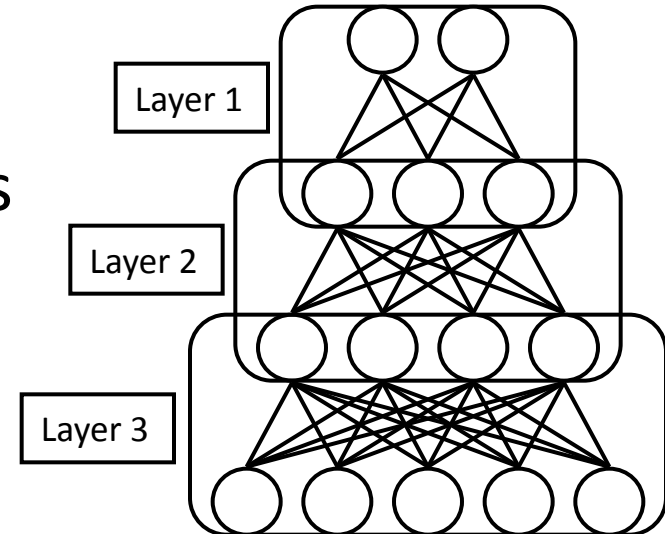


Training

- Contrastive divergence method
 - Activate network normally
 - Activate network with inputs “clamped”
 - Adjust weights to make normal activation behave more like clamped activation

Deep Belief Networks

- Use individual RBMs as layers in a larger network
- Hidden layer of one RBM forms input layer of another
- If single RBMs learn features about data, DBMs learn features about features



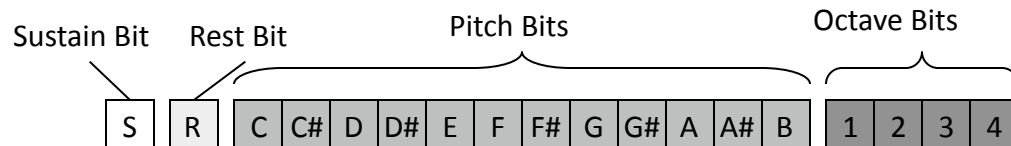
Encoding Scheme

- Requirements
 - Binary encoding
 - Music must be encoded in a string of standard length
 - Each note must be the same “distance” from every other note

Encoding Scheme



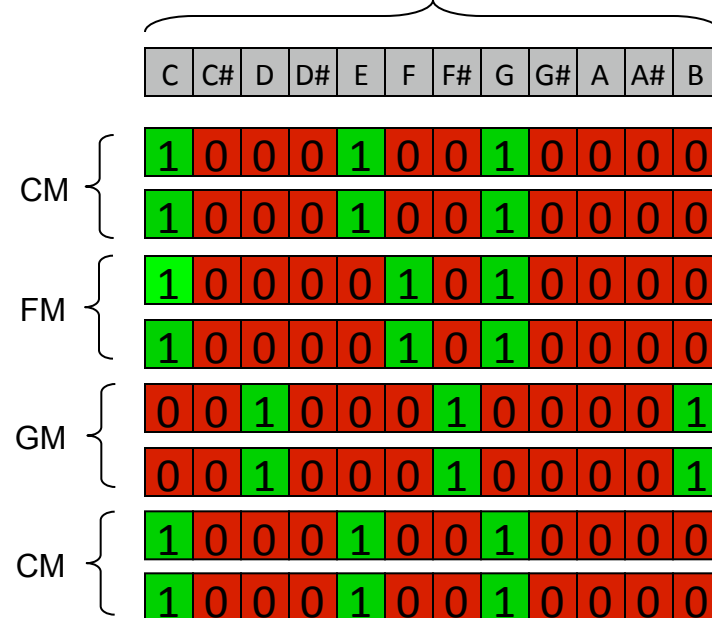
- Break melody into beat subdivisions
- Each subdivision contains 18 bits
 - 1 Sustain Bit
 - 1 Rest Bit
 - 12 Pitch Bits
 - 4 Octave Bits



Chord Encoding Scheme



Chord Bits



Initial Dataset

- Children's songs
 - 2 measures
 - 8th note resolution
 - Simple chords
 - 14 melodies
- Generated similar songs

Main Dataset

- Jazz licks
 - 4 measures
 - 12 beat subdivisions
 - (32nd triplet note resolution)
 - ii-V⁷-I-VI⁷ chord progression
 - 100+ licks
 - Transcribed
 - Handwritten

Windowing

- Rather than training on entire piece at once, break music into “windows”
- Start with the first measure of a lick and gradually move window forward
- Allows learning/generating arbitrary length music sequences with a fixed size network

Windowing



A musical staff in treble clef with a common time signature (C). The staff contains a sequence of notes: a quarter rest, followed by a series of eighth and quarter notes. A blue rounded rectangle highlights the first five notes of this sequence, which correspond to the Dm⁹ and G¹³ chords. The notes are: D4 (quarter), F#4 (quarter), G4 (quarter), F#4 (quarter), and E4 (quarter). Above the staff, four chord symbols are positioned: Dm⁹ above the first note, G¹³ above the fifth note, C^{maj7} above the eighth note, and A^{7/#5/#9} above the thirteenth note. The rest of the staff contains a continuation of the melodic line with various rhythmic values.

Windowing



A musical staff in treble clef with a common time signature (C). The staff contains a sequence of notes: a whole rest, followed by a series of eighth and quarter notes. A blue rounded rectangle highlights a window of six notes: F#4, G4, A4, B4, C5, and B4. Above the staff, four chord symbols are positioned: Dm⁹ above the first note, G¹³ above the second note, C^{maj7} above the third note, and A^{7/#5/#9} above the fourth note.

Dm⁹ G¹³ C^{maj7} A^{7/#5/#9}

Windowing



A musical staff in treble clef with a common time signature (C). The staff contains a sequence of notes: a whole rest, followed by a quarter note G4 with a sharp sign, a quarter note A4 with a natural sign, a quarter note B4 with a natural sign, a quarter note C5, a quarter note D5, a quarter note E5, a quarter note F5, a quarter note G5, a quarter note F5, a quarter note E5, a quarter note D5, a quarter note C5, a quarter note B4, a quarter note A4, a quarter note G4, and a quarter note F4. Above the staff, four chord symbols are positioned: Dm^9 above the first measure, G^{13} above the second measure, C^{maj7} above the third measure, and $A^{7/\#5/\#9}$ above the fourth measure. A light blue rounded rectangular window highlights the notes in the second measure, which correspond to the G^{13} chord.

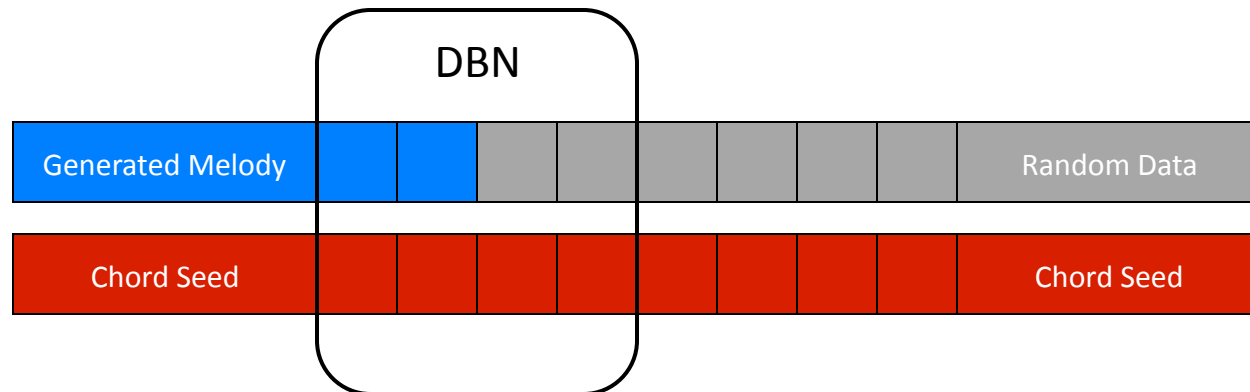
Generating New Melodies

- Need to specify a chord sequence over which to generate a new melody.
- Chord bits are “clamped” during generation so that they can influence the melody being generated without changing themselves

Generating New Melodies

- Use a windowing strategy analogous to our windowed training method
- As each successive beat is generated the whole melody and chord sequence shifts forward to make room for the next beat

Generating New Melodies



Results

- Rhythmically stable
- Respects chord tones
- Occasional color tones
- Very few foreign tones

Results

- Trained on transpositions as well
 - Generated music following key of given chord progression
 - Succeeded with up to four transpositions

Example Training Licks

Style: swing
Dm9

1 2 3 4

G13 CM7 A7#5#9

5 6 7 8

Dm9 G13 CM7 A7#5#9

9 10 11 12

Dm9 G13 CM7 A7#5#9

13 14 15 16



Example Generated Licks

5b1k cab92

Dm9 G13 CM7 A7b5#9

2 3 4

5 Dm9 G13 CM7 A7b5#9

6 7 8

9 Dm9 G13 CM7 A7b5#9

10 11 12

13 Dm9 G13 CM7 A7b5#9

14 15 16

17 Dm9 G13 CM7 A7b5#9

18 19 20

21 Dm9 G13 CM7 A7b5#9

22 23 24

25 Dm9 G13 CM7 A7b5#9

26 27 28

29 Dm9 G13 CM7 A7b5#9

30 31 32

2 2 2 2



More Generated Examples

3 **Dm9** 3 3 3 4 **G13**

5 **Dm9** 3 3 3 6 **G13** 3

7 **Dm9** 3 3 3 8 **G13** 3

Random Music

Style: no-style-but-swing

The image displays a musical score for four staves, likely representing different instruments or voices. The score is written in 4/4 time and features a variety of rhythmic patterns, including numerous triplets. The notes are color-coded: red for the upper staff, green for the second and fourth staves, and blue for the third staff. The first staff begins with a treble clef and a common time signature 'C'. The second staff starts with a bass clef and a '2' below it. The third and fourth staves also begin with a common time signature 'C' and are marked with '3' below them. The music is characterized by frequent triplet markings and a mix of eighth and sixteenth notes.



Future Work – Repeated Notes

- Our machines produced disproportionate numbers of repeated notes
- Can sound static or too immobile for jazz

Future Work – Repeated Notes

- Possible solution: post processing
 - Merge repeated notes together
 - Results in a smoother output, but starts to cross line of unsupervised learning
- Ideally, machine should avoid repeated notes in the first place

Future Work – Training Algorithm

- Slow
 - Optimization
 - Parallelization
 - Adaptive termination
- Sensitive to training presentation order
 - Randomize training inputs

Future Work – Chord Inference

- We believe our work naturally lends itself to the open problem of inferring unknown chords for a melody
 - Currently we provide a chord seed to generate a melody.
 - If we instead provide a melody as input, we could determine which chords fit that melody

Conclusion

- Unsupervised learning algorithm
- Based on probabilistic neural network theory
- Able to create novel jazz licks based on an existing corpus
- Minimal assumptions about musical knowledge