Matrosov M., Strijov V.V. Short-Term Forecasting of Musical Compositions Using Chord Sequences // Conference of the International Federation of Operational Research Societies, 2014 : 229.

2 - Scheduling Set-Up Operations in a Multi-Machine Environment when only One Set-Up Operator is Present
Daniel Schnitzler, Dirk Briskorn

There are a limited number of machines which have pre-assigned tasks. The tasks on a machine have to be processed in a given sequence. For each task, the machine has to be set up. Only one machine can be set up at a given time. Different goals are pursued (e.g., reduce makespan). Since standard solvers are only able to tackle small problems, a genetic algorithm and a tabu search were developed, which can solve problems with up to 100 machines and 1000 tasks. Different variants of the metaheuristics were tested with the help of random instances and instances from which the solution is known.

3 - Scheduling Part Feeding from Line-Integrated Supermarkets to Mixed-Model Assembly Lines
Simon Emde, Nils Boysen

Line-integrated supermarkets constitute a novel in-house parts logistics concept for feeding mixed-model assembly lines. In this context, supermarkets are decentralized logistics areas located directly in each station. Here, parts are withdrawn from their containers by a dedicated logistics worker and sorted just-in-sequence (JIS) into a JIS-bin. From this bin, assembly workers fetch the parts required by the current workpiece and mount them during the respective production cycle. This presentation treats the scheduling of the part supply processes within line-integrated supermarkets.

4 - Efficient Task Scheduling in Long-Term Care Facilities
Alexander Lieder, Dennis Moeke, Raik Stolletz, Ger Koole

Care workers in nursing homes are responsible for providing services to clients and cause the largest share of operational costs. In order to deliver high-quality service, it is important to assign each task to a qualified care worker and to a point in time according to the client's preferences. We present a dynamic programming approach that generates optimal task schedules. Using data from practice, we evaluate the runtime performance of this approach. A sensitivity analysis shows effects of optimal task schedules on the required workforce.

■ FA-15

Friday, 8:30-10:00 - Room 125

Experimental Research in Management Accounting and Management Control 2

Stream: Experimental Perspectives and Challenges in Management Accounting and Management Control
Invited session
Chair: *Stephan Leitner*

1 - Heuristic Methods for Picking Items for Experimental Sets
Rachel Bunder, Natashia Boland, Andrew Heathcote

Psychologists are often required to create sets of items to be used in experiments. Such sets are used to test how factors affect some situation, e.g., to see how humans respond to short words compared to long words. These sets must 'match', i.e., be as similar as possible, on all other attributes that could affect response. Previously, we have explored definitions of similarity for experimental data sets and have developed a MIP to solve this problem, which struggles when solving larger problems. We explore a variety of heuristic methods, comparing the results to existing metaheuristics.

2 - The Impact of Visualizing Causal Relations on Dynamic Decision Making
Michael Leyer, Jürgen Strohhecker

According to natural decision models, good decisions are mainly dependent on understanding the consequences of chosen options. Thus, receiving information on causal relations between options and results should be helpful. Using a capacity management simulator, we conducted laboratory experiments with two levels of complexity in which participants had to make decisions repeatedly. Results are showing not only key performance indicators on the user interface but also visualizing causal relations between them leads to better decisions. The results are stronger in the more complex situation.

3 - Impact of Information Overload on Escalation of Commitment
Peter Rötzel

Escalation of commitment explains why decision-makers are tempted to reinvest further resources in a losing course of action. While previous studies focus on the quality of information, there is a lack of research on how different information quantities affect escalation of commitment. Our study shows how information overload influences escalation of commitment and how information overload interacts with the decision-maker's earlier decisions. Our results indicate that decision-makers who face information overload increase their reinvestment even when the decision consequences are positive.

■ FA-16

Friday, 8:30-10:00 - Room 127

Pattern Recognition

Stream: Intelligent Optimization in Machine Learning and Data Analysis
Invited session
Chair: *Ivan Reyer*

1 - On Fingerprint Image Compression Method based on NMF
Congying Han, Tiande Guo

A new method for fingerprint compression is proposed. A general model that can be used to describe many existing algorithms, such as PCA, SVD and NMF is given. Based on the model, a modified NMF algorithm is used to train and compress images of fingerprint. A large number of tests show the new algorithm is valid for fingerprint compression. In particular, the method has a good performance for fingerprint with small size.

2 - Short-Term Forecasting of Musical Compositions Using Chord Sequences
Mikhail Matrosov, Vadim Strijov

The objective is to predict a sequence of chords. It is treated as multivariate time series of discrete values. A chord is represented as an array of half-tone sounds within one octave. We utilize a classifier based on probability distributions over chord sequences that are estimated both on a big training set and some revealed part of the forecasted melody. It shows robust forecasting on a set of 50 000 midi files. The novelty is model selection algorithm and invariant representation of chords. The same technique can be used to predict or synthesize various types of discrete time series.

3 - Parametric Shape Descriptor based on a Scalable Boundary-Skeleton Model
Ivan Reyer, Ksenia Zhukova

A parametric shape descriptor containing the set of convex vertices of a polygonal figure approximating the raster image and estimations of significance for curvature features corresponding to the vertices is suggested. The significance estimations are calculated with use of a family of boundary-skeleton shape models generated by the polygonal figure. Applications of the shape descriptor to face profile segmentation and content based image retrieval are presented.

4 - Customer Loyalty in Internet Service Provider Companies
Ilayda Ulku, Mehmet Yahya Durak, Fadime Üney-Yüksektepe

Internet is a basic standart of life and there are numerous service providers to make people safe, they try to service best quality and performance. Due to competition, providers try to prevent losing customer. In this research, a questionnaire is applied to get and analyze customer information, behavior and loyalty status of the churn possibility. This study deals with existing data mining algorithms to introduce the important factors for the churn prediction.

Short-term forecasting of musical compositions using chords sequences

Mikhail Matrosov[1,2,3],
supervised by Dr. Vadim Strijov[1,2,3],
consultant Anton Matrosov[1]

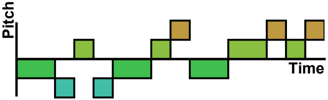
[1]Moscow Institute of Physics and Technology
[2]Skolkovo Institute of Science and Technology
[3]Computing Centre of the Russian Academy of Science

International Federation Of Operational Research Societies
July 2014, Barcelona, Spain

Research goal

Predict the next element in a sequence of chords of a special kind, not regarging temporal components (arpeggio, duration, pauses). The novelty is more accurate (with more details and consequently more data for each chord) representation of music sequence.

, not

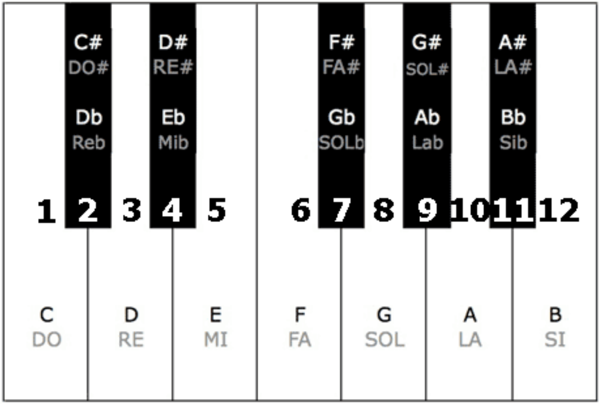


A sequence of simple tones

Teaser: 92.5% Hamming similiarity (58% chord-wise) between prediction and original melody on testing set.

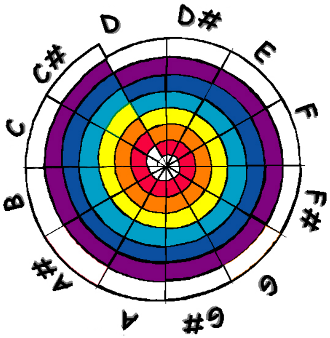
Octave

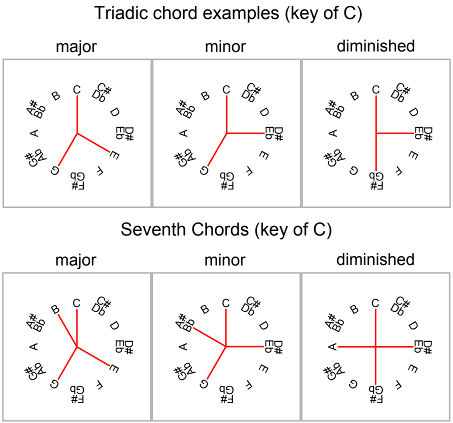
Octave consists of 12 semitones. Playing a composition, one can shift all notes into just one octave, and melody will sound almost the same (at least it will remain recognizable by an expert). So we keep chords within just one octave.



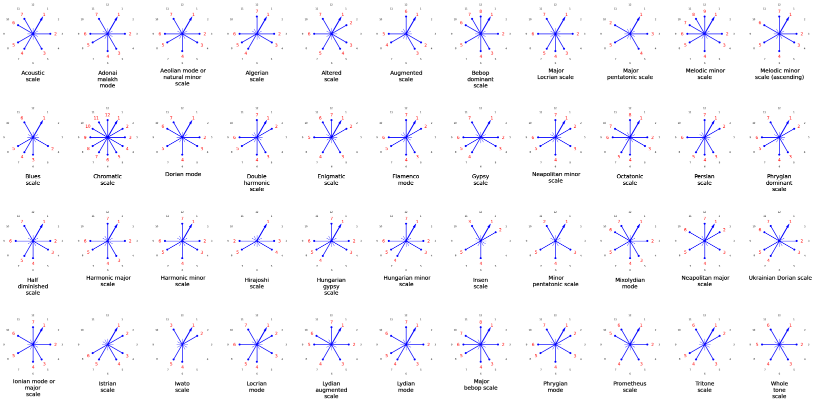
Chords

Each **chord** consists of 1 to 12 simultaneously sounding tones. Chord can be shifted several tones up or down. It can also be drawn on a circle, where rotation is changing pitch up or down.





Here are some more chords explained. Overall, there are 351 possible pitch constellations, or **strums**.



Each of these can be played in 12 different keys (transpositions), except for several simmetric cases (that are pretty rare in real music). That gives us $2^{12} - 1 = 4095$ possible chords.

Representation of a chord

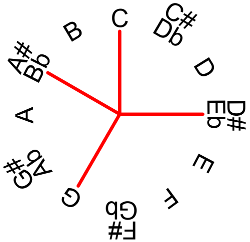
A melody can be represented as a sequence of chords. Each chord is an integer between 1 and $2^{12} - 1 = 4095$:

$$\underset{\text{melody}}{\mathbf{c}} = \underset{\text{sequence}}{\{c_i\}}, \ c_i \in \mathbf{C},$$

$\mathbf{C} = \{1, 2, 3, \ldots, 4095\}$ — space of **chords**.
Each chord has its base form (strum $s \in \mathbf{S}$) and key ($z \in \mathbf{Z}_{12}$). So the whole melody can be represented as a sequence of pairs (s, z):

$$\underset{\text{chord}}{c} = \underset{\text{strum}}{s} \times \underset{\text{key}}{z},$$

$$\underset{\text{melody}}{\mathbf{c}} = \underset{\text{pairs}}{\{(s, z)_i\}},$$



Multiplication stands for transposition.
$\mathbf{S}$ — set of unique **strums**, 351 elements.
$\mathbf{Z}_{12}$ — residue classes modulo 12.

Sequence of Elements

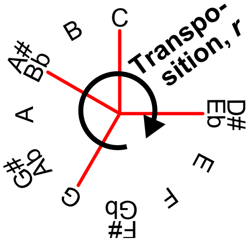
Melody can be transposed, so each key is relative to previous keys. That is why we use adjacent differences of keys:

$$r_i = z_i - z_{i-1} \mod 12, \; r_1 = z_1.$$

Pair (s, r) is called **element** and denoted as $x \in \mathbf{E}$.

$$\mathbf{C} = \mathbf{S} \times \mathbf{Z}_{12} = \mathbf{S} \times \mathbf{R}_{12} = \mathbf{E},$$

$\mathbf{C}$ — space of **chords**, $N = 4095$,
$\mathbf{S}$ — set of unique **strums**, $N = 351$,
$\mathbf{Z}_{12}$ — residue classes modulo 12, $N = 12$,
$\mathbf{R}_{12}$ — differences of Z_{12}, $N = 12$,
$\mathbf{E}$ — space of **elements**, that we predict.
Space is a set with an operation
of transposition (moving pitch up or down).



N-grams

N-gram is a contiguous sequence of N elements $x \in \mathbf{E}$ from a given sequence $\mathbf{x} = \{x_i\}$. N-gram of size 1 (unigram) is just one element $x \in \mathbf{E}$. For example:

$$\mathbf{x}_k^N = \{x_k, x_{k+1}, \dots, x_{k+N-1}\}.$$

In this work N-grams are used as features describing current point in the music sequence.

Dataset

50 000 random midi files were grabbed from the Internet. Each midi file was converted to a sequence of chords $\mathbf{c} = \{c_i\},\ c_i \in \mathbf{C}$ with the following steps:

- open midi file as a piano roll,
- remove percussion part,
- quantize time with rate $2 \cdot tempo$,
- strip octave number ($new\ pitch = pitch \mod 12$).

Average midi file contains sequence of **600 chords**, that gives **30 million** chords overall.
Melody is a sequence of elements (index is time): $\mathbf{x} = \{x_i\},\ x_i \in \mathbf{E}$.
$\mathbb{X}$ is a set of melodies: $\mathbb{X} = \{\mathbf{x}_j\}$.

Train/test dataset division

For evaluation purposes full dataset (50 000 melodies) was being splitted several times in two pieces of different size. Each time splitting was performed on a random basis — from the dataset was selected a subset (without returns) of requested size M.

$$\mathbb{X}_{training} \subset \mathbb{X}_0,$$

$$|\mathbb{X}_{training}| = M.$$

To test the algorithm we use the rest of the dataset:

$$\mathbb{X}_{testing} = \mathbb{X}_0 \setminus \mathbb{X}_{training}.$$

Prediction function

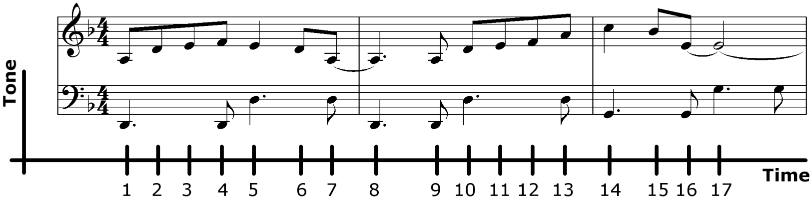
Prediction is made using weighted sum of several classifiers:

$$x_{k+1} = f(\mathbb{X}, \{x_1, \dots, x_k\}, \underbrace{\mathbf{w}}_{=\{u_i, v_i\}}) = \arg\max_{e \in \mathbf{E}} \sum_i (A_{ei} u_i + B_{ei} v_i),$$

$$A_{ei} \propto N\left(\underbrace{\{x_{k-i+1}, \dots, x_k, e\}}_{\text{(i+1)-gram}} \text{ in } \underbrace{\mathbb{X}}_{\text{Training set}}\right),$$

$$B_{ei} \propto N\left(\underbrace{\{x_{k-i+1}, \dots, x_k, e\}}_{\text{(i+1)-gram}} \text{ in } \underbrace{\{x_1, \dots, x_k\}}_{\text{Part of melody before k+1}}\right),$$

where "$\propto$" means that A_{*i} and B_{*i} are L1-normalized, $x_k \in \mathbf{E}$,
$N(g \text{ in } \textit{dataset})$ — number of n-grams $g \in \mathbf{E}^n$ in *dataset*,
$\mathbf{w} = \{u_i, v_i\} \in \mathbb{R}^{2K}$ — vector of model parameters (weights).



Function f is a classifier that predicts the next element. Then error function is (k denotes time interval)

$$S(\mathbf{w}, \mathbb{X}) = \sum_{\mathbf{x} \in \mathbb{X}} \sum_{k=K}^{N_{\mathbf{x}}-1} \left[x_{k+1} \neq f(\mathbb{X}, \underbrace{\{x_1, \ldots, x_k\}}_{\text{Prev. part of the melody}}, \mathbf{w}) \right].$$

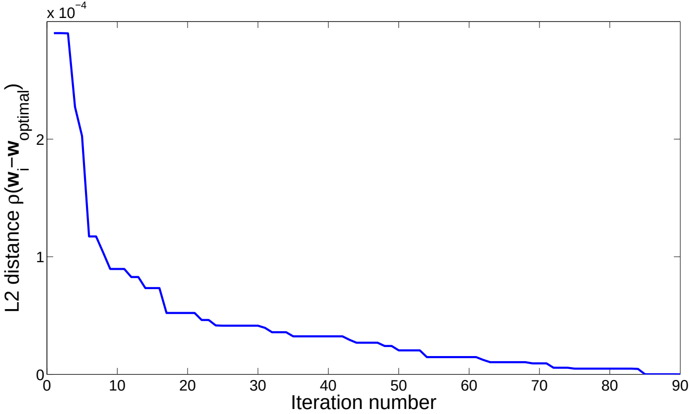
Brackets stand for 1 if the statement inside is true and 0 if false. $\mathbb{X}$ is a set of melodies, $\mathbf{w} \in \mathbb{R}^{2K}$ is vector of parameters.

Problem statement for optimization

For a training dataset $\mathbb{X}$ we would like to find vector $\mathbf{w}$ of algorithm parameters (weights), that minimizes error function:

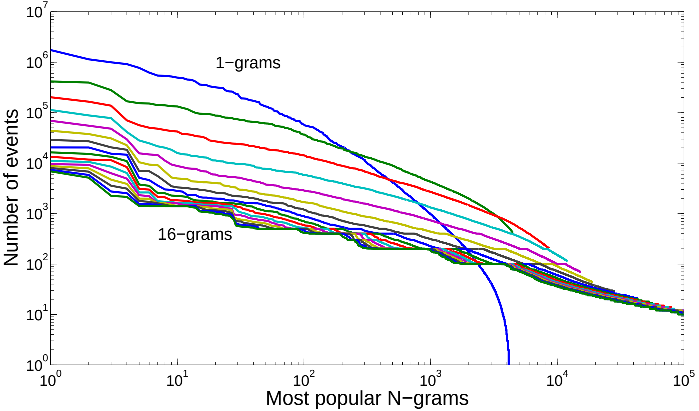
$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathbb{R}^{2K}} S(\mathbf{w}, \mathbb{X}).$$

One evaluation of predicting and error function can take 100 hours (on my laptop). Therefore we make **small steps** for just a small part (bunch) of training set. One bunch is typically **100 melodies** (random subset) comparing to 10 000 usual dataset and it can be evaluated **much faster** because it fits into RAM.



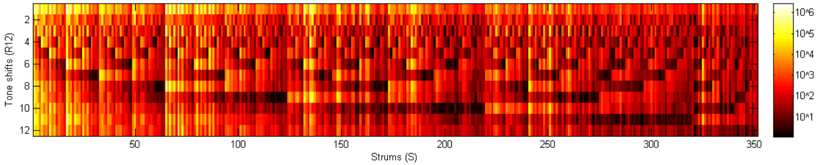
Interesting fact

Probability distributions of N-grams with different N. Number of events is a number of occurences found in a set of 50 000 midi files. Slope coefficient is about -0.6, distribution is similiar to distribution of words in a natural language (Zipf's law).



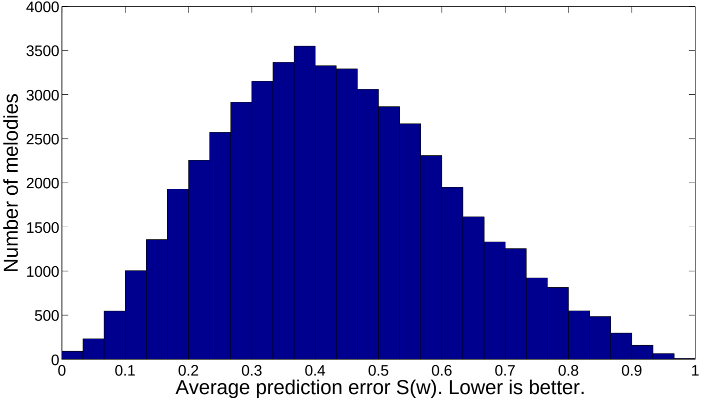
Visualization

Heat map of distribution probability of elements **E**. Number of events is a number of occurences found in a set of 50 000 midi files. Order of strums (horizontal axis) is arbitrary — result of representing a chord as a pair (s, r).



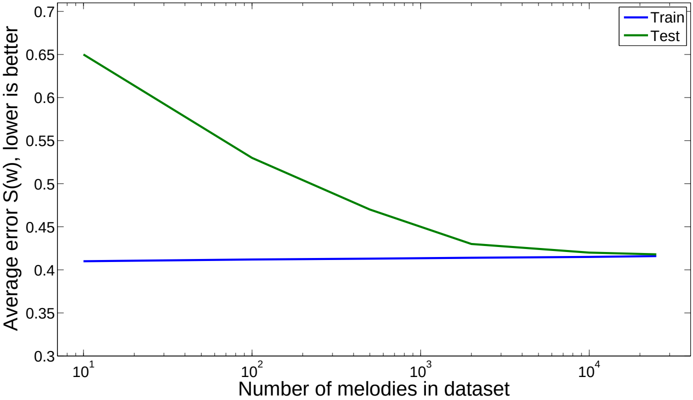
Forecasting quality

Number of parameters is 16, training set is 50 000 melodies. Average prediction error is **0.42** (meaning 58% of successfully predicted elements $e \in \mathbf{E}$). There are also melodies that were forecasted on 100%, as well as melodies forecasted poorly (<5%).



Training data size

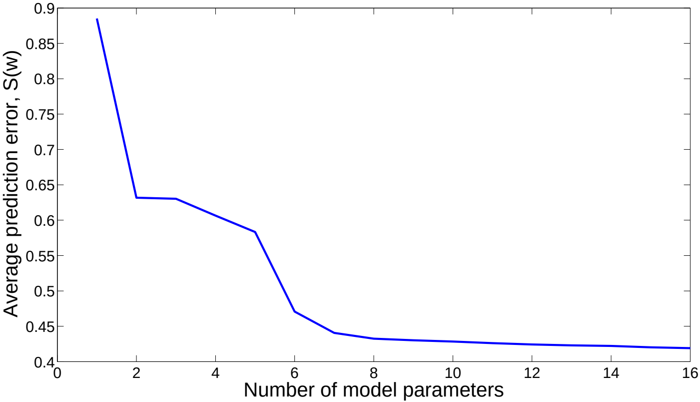
$$S(\mathbf{w}, \mathbb{X}) = \sum_{\mathbf{x}\in\mathbb{X}} \sum_{k=K}^{N_\mathbf{x}-1} \left[x_{k+1} \neq f(\mathbb{X}, \underbrace{\{x_1, \dots, x_k\}}_{\text{Prev. part of the melody}}, \mathbf{w}) \right].$$



Number of parameters

$$S(\mathbf{w}, \mathbb{X}) = \sum_{\mathbf{x}\in\mathbb{X}} \sum_{k=K}^{N_{\mathbf{x}}-1} \left[x_{k+1} \neq f(\mathbb{X}, \underbrace{\{x_1, \dots, x_k\}}_{\text{Prev. part of the melody}}, \mathbf{w}) \right].$$

Error function vs number of parameters K ($\mathbf{w} \in \mathbb{R}^{2K}$), test set:



Existing methods

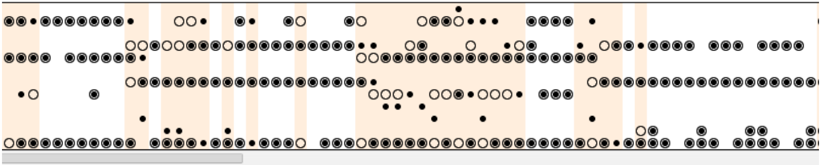
Quality	Mozer[1]	Conklin[2]	Proposed
Main idea	Neural network	Music patters	Bayes classifiers
Chords	—	40%	**58.0%**
Pitches	93%	**95%**	92.5%
Durations	**90%**	75%	—
Datasize	20	4500	50 000

[1] *Neural network music composition by prediction* — M. Mozer, Connection Science, 1994.

[2] *Multiple viewpoint systems for music prediction* — D. Conklin, I. Witten, Journal of New Music Research, 1995, rev. 2002.

Results

Forecasting example: circles represent the truth tones, dots — predicted tones, errors are highlighted, horizontal axis is time.



- The optimal model complexity (max combination length) is 8, thought the more, the better.
- Number of songs in the training set should be at least 1000.
- Forecasting quality is 58% (chord-wise, 0.024% for a random guess), Hamming distance is 0.075 (meaning 92.5% tone matches comparing to 50% for a random guess).