

Kent Academic Repository

Full text document (pdf)

Citation for published version

Jordanous, Anna and Smaill, Alan (2008) Artificially intelligent accompaniment using Hidden Markov Models to model musical structure. In: Conference on Interdisciplinary Musicology, 2008, Thessaloniki, Greece.

DOI

Link to record in KAR

<https://kar.kent.ac.uk/42368/>

Document Version

UNSPECIFIED

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

Artificially intelligent accompaniment using Hidden Markov Models to model musical structure

Anna Jordanous

Music Informatics, Department of Informatics, University of Sussex, UK
a.k.jordanous at sussex.ac.uk - <http://www.informatics.sussex.ac.uk/users/akj20>

Alan Smail

School of Informatics, University of Edinburgh, UK
smaill at inf.ed.ac.uk - http://www.inf.ed.ac.uk/people/staff/Alan_Smail.html

Proceedings of the fourth Conference on Interdisciplinary Musicology (CIM08)
Thessaloniki, Greece, 3-6 July 2008, <http://web.auth.gr/cim08/>

Background in Music Performance and Accompaniment. Musical accompanists may not always be available during practice, or the available accompanist may not have the technical ability necessary. As a solution to this problem, many musicians practise with pre-recorded accompaniment. Such an accompaniment is fixed and does not interact with the musician's playing: the musician must adapt their performance to match the recording. It is more natural for the musician if the accompaniment adapts to fit the performer. To synchronise accompaniment with soloist, an accompanist should be able to follow the musician through the score as they play. Complications arise if the performer deviates from the score: either intentionally, by adding their own musical interpretation, or accidentally, by making performance errors. The accompanist should be able to adjust to such behaviour.

Background in Computing, mathematics and statistics (in musicology). There are several ways to give mathematical models of statistical information associated with discrete linear sequences of observations. Hidden Markov Models (HMMs) work by supposing that the observations depend statistically on some hidden states of the system, and on the most recent hidden states and observations (Rabiner (1989)). The associated statistical information can be learned algorithmically, or estimated otherwise; such a system is then able to generate new observation sequences that exhibit the same statistical patterns. This approach has proved effective in capturing local properties of sequential data in many areas, e.g. biology (Durbin et al. (1998)), as well as in musicological analysis.

Aims. This work investigates how an intelligent artificial musical system can follow a human musician through the performance of a piece (perform *score following*) using a Hidden Markov Model of the piece's musical structure. The system interacts with the human musician in real-time and provides appropriate musical accompaniment.

Main Contribution. Prior research using HMMs in score following has concentrated on modelling note onsets, durations and offsets (e.g. Cano et al 1999), or on modelling events in the score: patterns of notes that have been selected as significant to look out for (e.g. Orio et al 2003). This work concentrates on modelling the musical structure of a piece by using HMM states to represent individual beats of the piece, an approach to score following which to the best of the authors' knowledge has not been tried before. Here we were influenced by recent advances in beat tracking (Gouyon and Dixon 2005). Having successfully implemented this representation, the performances of the resulting artificial accompanists were evaluated by human testers and by using objective criteria based on that used at the Music Information Retrieval and Exchange conference in 2006. Accompaniment accuracy was measured at a total precision level of 60.89% and piecewise precision of 54.05%, comparable to systems tested at MIREX 2006.

Implications. Using HMMs considerably simplified implementation of our artificial accompanists. We did not need to consider how the artificial accompanist tracked the soloist through the score, beyond modelling the musical structure using an HMM. This compares favourably to other score following methods that have been tried (e.g. Vercoe 1984, Dannenberg 1989), where a score following algorithm must be specifically encoded.

An unforeseen but fascinating result of this experimentation was the observed importance of co-operative feedback and communication in the performer/accompanist scenario in real-life. To date, score following research has concentrated on the artificial accompanist following the soloist, however our results suggest it would be worthwhile to pursue future research that focus more on co-operation between soloist and accompanist.

Consider a flautist who is performing a solo piece at a concert, with a pianist providing accompaniment. The piano accompanist listens to what the flautist is playing, to ensure their accompaniment matches the flautist. The flautist's performance may occasionally deviate from what is written in

the score. In these cases the piano player adjusts their accompaniment accordingly.

Score following is the process where a musician follows another musician's playing of a musical piece, by tracking their progress through the score of that piece. The term is

most commonly used in the context of computer-generated accompaniment, where one or more of the musicians involved are artificial rather than human.

Early attempts to implement score following centred around dynamic programming and pattern matching (Vercoe (1984), Dannenberg (1984)). Probabilistic methods were first attempted by Grubb and Dannenberg (1998). This work paved the way for the use of Hidden Markov Models (HMM), a stochastic modelling technique, which has emerged as a promising way of implementing score following (Orio and Dechelle (2001), Raphael (2001), Schwarz et al (2004), Pardo and Birmingham (2006)).

During this research, artificially intelligent accompaniment systems were developed. These artificial accompanists used a Hidden Markov Model representation of the musical structure of a piece of music, to follow the soloist's progress through the musical score and provide accompaniment in real-time.

We chose to use a slightly different approach in fitting Hidden Markov Models to the music being performed: modelling the music beat by beat, as opposed to identifying significant events in the score (e.g. Orio et al 2003), or using HMMs to process the incoming audio signal (e.g. Cano et al 1999). To the best of our knowledge, this use of HMMs to model musical structure by beat rather than by note or musical event has not been attempted before in score following research.

The more advanced systems developed during our research incorporated beat tracking, complex accompaniment and relative score positioning.

Three musical pieces of varying complexity and length were programmed into the different versions of the artificial accompanists. Testers of varying musical ability and experience evaluated the performance of each artificial accompanist subjectively. Each artificial accompanist was also tested objectively by criteria that was used to evaluate artificial accompanists at the Music Information Retrieval Evaluation eXchange conference in 2006; hence some general comparisons could be made between

this work's score following systems and alternative score following systems.

The purpose of this work was to test the practicality and efficiency of our HMM representation of musical structure by beat, specifically for an artificially intelligent accompaniment system. This paper presents our findings, highlighting some interesting observations that arose during this research.

Accompaniment issues for musicians

Think back to the flautist performing alongside a piano accompanist. There are several reasons why the flautist may not perform the piece exactly as written. They may make mistakes: missing some notes out, misplaying others or adding extra notes. In addition the flautist should have the freedom to add musical embellishments that do not exist in the original score, without this disrupting the accompaniment. The accompanist should adjust their playing, according to any such deviations from the written score by the flautist. Providing musical accompaniment, then, is not necessarily a straightforward process.

Outside of performance, it is also useful for a musician to have access to accompaniment during practice. The musician can learn how the accompaniment sounds. From this they can derive valuable assistance for future performance. As example, the musician would be aware of the underlying harmony provided by the accompaniment, and of any musical cues they could use.

Accompanists may not always be available when needed for practice or performance. A related problem is that the accompanists available may not have sufficient technical ability to provide adequate accompaniment.

One possible solution to these problems is to use accompaniment that has been generated automatically by a computer or recording. Many musicians practise playing over recorded or computer generated accompaniment where the accompaniment is static, i.e. it will not change from one performance to another. This means, though, that the musician may need to adapt their performance to match the recording.

It is more natural for the musician if the accompaniment adapts to fit the performer. Raphael (2001) describes this as moving from "music minus one" to "music plus one". To dynamically synchronise the accompaniment with the performance by the musician, it would be necessary for the accompanist to track the performer in some way through the score of the piece as they play. This may become complicated if the performer deviates from the score.

Hidden Markov Models

Hidden Markov Models (HMM) are a stochastic modelling tool, popular in a variety of domains from speech processing (Rabiner (1989) to biological sequence matching (Durbin et al, (1998)).

Real-world systems that produce some kind of observable signal can be modelled with HMMs. In particular this includes systems that operate non-deterministically: systems whose behaviour cannot be predicted exactly by using algorithmic rules or formulae.

Probabilities are used in the HMM to represent the system's observable behaviour and to represent internal (hidden) facets of the system. The HMM can then be used to process these observable signals to explain the system's behaviour and make probability-based estimates about future behaviour.

As Rabiner describes in his comprehensive tutorial (Rabiner (1989)), a system modelled with an HMM can be considered to be in one of a finite number of states at any given time. We can gain information about what state the system is currently in by examining recent outputs from the system ('observations'). The actual states themselves can not be observed, just the sequence of observations that result from the system passing through those states. The observed output can be interpreted as being "a probabilistic function of [the system being in] the state" (Rabiner (1989), p. 258).

The relationship between individual states and observations is not a functional relationship but a many-to-many relationship; one observation may be produced by many system states, and in turn there may be more than one possible observation should the

system be in a given state. Durbin (1989) highlights the fundamental difference between Hidden Markov Models and Markov chains: with an HMM, you cannot gauge what state the sequence is in purely from the current observation in isolation. There is not a one-to-one correlation between states and observations with HMMs, though there is with Markov chains.

Score Following using HMMs

A musical score is divided up into a sequence of musical events (for example where one note is considered as one modellable musical event).

The artificial accompanist is given a Hidden Markov Model that represents these musical events, and uses an algorithm such as the Viterbi algorithm to estimate what state the performer is most likely to be in at that time, i.e. which musical event in the score the performer is currently playing.

The aim is to find the most probable state sequence that generated the given sequence of observations (notes played by the soloist).

This work uses the Viterbi algorithm is used to find out which state the soloist is most likely to be in (given observations of recent notes played by the soloist).

Implemented in the traditional fashion, this algorithm finds the globally optimum path through the Hidden Markov Model states to the most probable current state, using the history of observations seen. However in score modelling we instead require a locally optimal path to the current point. This is because we are interested in the correct accompaniment playing at the right time, even if the resulting path through the music overall is not the most probable path when the performance is viewed as a whole.

To fit an HMM to a piece, events in the performance (for example rests, notes, trills, chords, and so on) are modelled by HMM states. The notes played by the soloist form the observations which the Viterbi algorithm uses to track the performer's progression through the score.

Whilst the model encapsulates the score of the music, it must also allow for cases when

the performer deviates from the score. Inspired by the approach taken by Orio and Dechelle (2001), we model each event in the piece in parallel with both a *normal* state and a *ghost* state. Normal states represent the state that the soloist is in if they are playing the piece as written in the score. Ghost states represent the state reached by the soloist if they have deviated from the score at that point. Figure 1 represents different types of transitions through normal and ghost states, corresponding to different performances by the soloist.

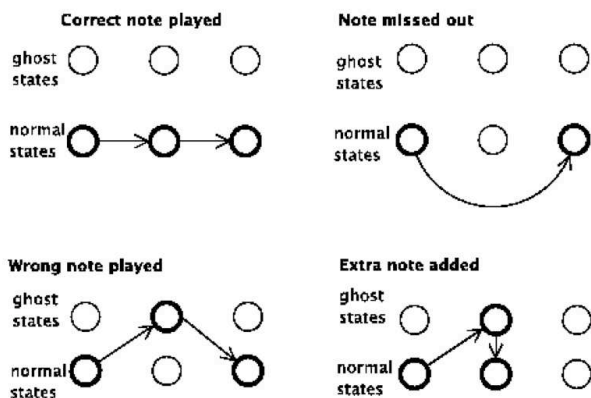


Figure 1. Transitions between *normal* and *ghost* Hidden Markov Model states in various performance scenarios.

A different approach to using HMMs: modelling the musical structure by beat

The work presented here places an emphasis on modelling the musical structure of a piece by using HMM states to represent individual beats of the piece, an approach to score following which to the best of the authors' knowledge has not been tried before. Here we were influenced by recent advances in beat tracking (Gouyon and Dixon 2005).

If this approach works successfully, then less reliance is placed on the programmer to identify key events, as in the approach where HMM states model important events in the score (Orio et al 2003 discuss a number of score following systems with this approach).

Developing artificial accompanists

The artificial accompanists were developed in Max/MSP, a programmable music processing

environment. MIDI input and output was through a Yamaha Clavinova. Using Max/MSP means that our system could in future be adapted to include signal processing in addition to MIDI input/output, if required.

We had difficulty finding a suitable Max/MSP implementation of Hidden Markov Models (HMM). Consequently the artificial accompanist system incorporated our implementation of a standard HMM model structure and the use of the Viterbi algorithm to analyse musical input from the soloist.

Extracts from three pieces were selected for performance by a human soloist and the artificial accompaniment system. These three pieces were each modelled by a Hidden Markov Model, such that the notes in the melody were treated as the observations connected to transitions between sequential *normal* states.

Melody 1 The first, from the traditional melody *Twinkle Twinkle Little Star*, was the most simple. It had a completely homophonic accompaniment, always moving in parallel with the soloist's melody.

Melody 2 An extract from Andrew Lloyd-Webber's *All I Ask Of You* offered the artificial accompanist task more variety of note lengths and a longer extract in total.

Two different accompaniments were arranged for this melody: an accompaniment with no movement independent of the soloist movement, and a second more complex accompaniment where the accompaniment moved between notes whilst the soloist remained holding one note.

Melody 3 *Danse Macabre* was selected specifically as a more challenging solo melody to track the soloist through. This is because it incorporates much repetition of note sequences, and some stylistic variation in note lengths

Evaluating the accompanists

The overall aim of a competent artificial accompanist should be to provide musical and accurate accompaniment, interacting with the performer in real time. The quality of an accompanist's performance in general is judged by how well it fits and enhances the playing of the soloist whom they are accompanying; the very nature of a good accompanist is that the audience is not aware of their playing except as an enhancement to the soloist's performance.

The performances of the artificial accompanists produced during this research were evaluated both objectively and subjectively.

The system was tested against measurable criteria originally constructed in 2006 by score following experts to test the latest research efforts (Cont and Schwarz, 2006).

As well as this testing, the artificial accompanists were tested and judged by musicians of varying musical ability and experience, so that they could give their opinions on the quality of accompaniment provided by the artificial accompanists.

Several versions of the artificial accompanists were tested, from simple versions with no beat tracking, to more advanced versions with beat tracking, use of more historical observations to track the soloist and amendments to the system for efficiency.

Methodology for quantitative evaluation

Taking testing criteria from the 2006 Music Information Retrieval Evaluation eXchange (MIREX) conference (Cont and Schwarz, 2006), our quantitative evaluation measured:

- Event Count (the number of musical events included in the played melody)
- False Positives (scored notes which are only recognised after a delay greater than 2000 milliseconds)
- Number of Notes Missed (this statistic is also inclusive of False Positive notes)
- Mean and Standard Deviation Offset (the difference between the soloist's note onset and the accompaniment note onset)
- Mean Latency (the difference between the detection time of the note being played by the soloist and the time the system has processed the audio so that it is ready to be matched to the score¹)
- Missed Note Percentage and False Positive Percentage

There were two additional measures we could use to compare our work overall with the artificial accompanists submitted at MIREX 2006:

- Total precision (percentage of correctly detected notes overall, i.e. results for all pieces, added together).

- Piecewise precision (mean of the percentage of correctly detected score notes for each piece by the artificial accompanist).

To make some quantitative measure of musicality and fluency of these performances, for later reference, in objective testing we included a rating, from 0 to 5, of how well we judged our artificial accompanist to have performed accompaniment during the test².

Five tests were carried out on each artificial accompanist. For each test, the artificial accompanist was presented with a specified melody from the soloist. Performance was measured using the above criteria.

1. Play the melody as scored, with no mistakes, tempo changes or embellishments
2. Play the melody with selected errors added
3. Play the melody with selected embellishments added
4. Play the melody as scored but with selected tempo adjustments made
5. Play the melody, making all the deviations from the score from tests 2, 3 and 4

Methodology for qualitative evaluation

In addition to testing the artificial accompanists against objective measurable criteria, the artificial accompanists that were developed in this work were evaluated by human musicians of different levels of musical competence and experience.

Four testers were presented with five versions of the artificial accompanist to test, in order of increasing complexity of artificial accompanist functionality and the piece.

For each piece, the testers were allowed up to five minutes to practice the solo melodies before adding the automatic accompaniment. This meant that they could pay more attention to the performance of the accompaniment rather than concentrating purely on playing the right note, but still made occasional unintended mistakes, especially for more complex melodies.

In each test, the testers were asked first to play the melody as correctly as they could, then to play it with different variations of mistakes, embellishments and tempo changes. They were asked to experiment with the system as they saw fit, using their

musical knowledge and imagination. We deliberately did not specify any errors or embellishments that the testers should make, to avoid influencing them.

The testers were asked to give comments during and after each piece, on how well they perceived the system to accompany them, focusing on how well it recovers from errors and embellishments that they added.

Results and discussion of evaluation

A comprehensive list of results and detailed discussion can be found in Jordanous (2007); here we present a summary.

It is pleasing to see in Table 1 that the artificial accompanists developed in this research compared favourably overall in performance to the two artificial accompanists analysed at MIREX 2006 (3,4). The weaker result on the piecewise precision is affected by the poor performances overall from the artificial accompanists with Melody 3. These comparisons, however, can only be made at a very general level, as our artificial accompanists were tested on different pieces to those presented at MIREX 2006.

Authors	Total Precision	Piecewise Precision
Arshia Cont and Diemo Schwarz (MIREX 2006)	82.90%	90.06%
Miller Puckette (MIREX 2006)	29.75%	69.74 %
This work	60.89%	54.04%

Table 1: Comparison of overall performance

During comparison, it was interesting to see a degree of variance in the accuracy of the MIREX 2006 artificial accompanists, depending on what piece is being played. This was also true for the different pieces that our artificial accompanist was tested on.

As expected, the artificial accompanists performed much better in accompanying the two simpler melodies than the more complex third melody. Both quantitative and qualitative testing provided evidence for this conclusion. Lower percentages were recorded in the Missed Note % and False Positive % measurements for the two simpler melodies, with average offset figures of 12-542ms as opposed to up to 982ms for the third melody.

Tester feedback was also more positive for the first two melodies. Testers judged the standard of accompaniment produced for the two simpler pieces to be superior to the third, with no noticeable latency issues.

An unsurprising observation was that the artificial accompanists incorporating some form of beat-tracking had higher latency measurements for receiving and processing the soloist's playing than for the simpler artificial accompanists (a difference of approximately 200ms in general). This is due to the extra processing involved.

In particular the artificial accompanists for Melody 1 and Melody 2 performed the accompaniment better than anticipated during Test 5. (This was the test where all the errors from the previous tests were combined into one playing.) Occasionally the test melody was almost unrecognisable from the original tune. A human accompanist would have had to apply some skill and concentration when accompanying a soloist who was making this number of deviations from the score. So the attempts made to accompany the soloist in Tests 5 were a very positive result of testing.

The artificial accompanist that used a history of four observations for the Viterbi algorithm gave a very accurate performance in the first test for Melody 3 (where the solo melody was performed correctly). It was also reasonably accurate in the second test (where selected errors were included during performance of the solo melody). This shows the improvements in accuracy possible if more information from the soloist is considered.

A criticism of this particular artificial accompanist, though, is that latency measurements associated with the more detailed calculations were considerably higher and this is reflected in the poorer ratings overall that the third version received for quality of accompaniment. In general, the higher the Average Offset or Average Latency recorded, the less musically accurate the artificial accompanist was judged to be.

There were very large figures (242 - 982ms) for the Average Offset (representing processing time) when testing with Melody 3. This was reflected in the performance, where the accompanist lagged behind the soloist (particularly in Tests 3 and 5).

However the overall accuracy measurements (missed notes and false positive) for some

tests on this melody were considerably higher than expected, given how the accompaniment was deemed to have performed by testers. The quantitative testing often revealed that the artificial accompanist was in fact locating the performer at the right point in the score, but not quickly enough. As the primary objective of a artificial accompanist must be to produce musically accurate accompaniment, this latency should be addressed in future work.

In evaluation, the testers generally judged the artificial accompanists as being able to detect changes in tempo rapidly, although further work is required to detect the magnitude of the change in tempo more objectively. This aspect of the artificial accompanist is related to the ability to track the performer accurately through the piece, so as the HMM probabilities are more accurately set, this aspect of the artificial accompanist works more competently.

There was however some variance between different testers which we believe is due to different playing styles. For example, in the third melody (which included a high proportion of staccato notes), our second tester achieved better results than the other testers as their interpretation of playing notes staccato was the closest to our interpretation used during development.

This highlights the usefulness of having several musicians' influence on the development of the musicality of the artificial accompanist (as is the case in real life; a human musician will usually benefit from a variety of different influences).

Our artificial accompanists in general performed better with musicians of lower rather than higher ability. They responded better to inconsistent tempos and errors, as opposed to decorative embellishments. This is probably partly due to a slight bias in the way we have set the HMM probabilities, towards recovering from errors rather than dealing with decorations and embellishments.

It is pleasing, however, to see that most of the artificial accompanists generally performed well in responding to tester errors of different types, and coped with note embellishments to a certain degree.

An unforeseen but fascinating result of the testers' experimentation with our artificial accompanist system was the emerging of the co-operative nature of this domain in real-life,

and the importance of feedback and communication between two musicians. Roger Dannenberg has commented on a similar finding in an ensemble situation (Dannenberg, 2000):

"Early on, Lorin [Grubb] and I were playing trios with the computer, making intentional errors to test the system. We found that if we deliberately diverged so as to be playing in two different places, the computer could not decide who to follow. Even if one of us played normally and the other made an abrupt departure from the normal tempo, the computer would not always follow the "normal" player. In a moment of inspiration, we realized that the computer did not consider itself to be a member of the ensemble. We changed that, and then the computer performed much more reasonably. Here is why this worked: When the computer became a first-class member of the ensemble and one of us diverged, there were still two members playing together normally, e.g. Lorin and the computer. The computer, hearing two members performing together, would ignore the third."

While the emphasis found in previous research, and in this work, has been on the artificial accompanist following the soloist, we believe that a design with more focus on co-operation between soloist and accompanist would be worth further investigation, having been neglected in score following research to date.

Achievements of this work

The artificial accompanists use an HMM representation of musical structure of a piece by beat, to follow a soloist through the performance of that piece.

As a result the artificial accompaniment system can produce musically acceptable accompaniment, even if the soloist's performance is occasionally inaccurate or embellished. The systems match the performer's interpretation in terms of the volume the soloist is playing at and have been judged as reasonably accurate in matching the performer's tempo.

The use of an HMM considerably simplified our implementation of score following. We did not have to give strong consideration to how the artificial accompanist tracked the soloist through the score, beyond implementing the HMM. So the performance of the resulting artificial accompanists is pleasing, and we feel

that the chosen HMM representation of the domain was justified.

With further experimentation as to the most appropriate settings for the HMM probabilities, and perhaps implementation of some automatic training for individual pieces or individual performance styles, better results should be possible for the artificial accompanists that did not perform so well.

These artificial accompanists used a simple implementation of beat tracking. This was made considerably more simple to implement due to our use of HMM states to represent beat structure, rather than notes in the score. The ease with which we could incorporate beat tracking in the accompanist proved the worth of our decision to use this novel approach in representing the musical structure, rather than following the representations described in previous work (e.g. Orio et al 2003, Cano et al 1999).

In testing, the beat tracking appeared to work quite well for the simple melodies, however there was a problem with more complex melodies. This was because of the reliance on the state to be located correctly in order to gauge note lengths, and therefore the expected distance between observations. An alternative implementation of beat tracking that could have been tried was to include the use of the previous tempo to work out roughly how many beats had passed between two note inputs, as opposed to relying on the HMM to have estimated the next state correctly in all occasions.

One area that needs further investigation is in the efficiency of the artificial accompanists. The more complex artificial accompanists in this research demonstrate how latency issues can severely disrupt the performance of the accompaniment by the artificial accompanist. Careful consideration needs to be made as to how to overcome the large calculation effort involved in larger scale score models (perhaps by using an alternative to the Viterbi algorithm or by optimising it further).

We note here that the concern with efficiency is not with the general use of an HMM structure, but specifically with the extraction of information from the HMM by calculations with the HMM probabilities.

So the findings of this research project are that an HMM is a good way to implement score following, but that the HMM probabilities need to be set carefully. Also

there are concerns about the efficiency of using the Viterbi algorithm to track the soloist through the score.

Future work

The following suggestions could all feasibly be added to our artificial accompanists⁵.

- Update the system to be able to process audio input/output as well as MIDI
- Automatically extracting the score and HMM structure (currently programmed by hand).
- See if machine learning techniques could help the artificial accompanist to learn a particular performer's common performances (by training the HMM)
- For assistance in teaching purposes: add a "tutor" that gives feedback to the performer on how they deviated from the score
- Add knowledge to the artificial accompanist that allows it to respond to musical cues and feedback from the soloist, to co-operate with the soloist in performance

The last of these suggestions is possibly the most intriguing. Our testers remarked on how the accompanist seemed to follow the tester too closely. In contrast, they would expect a human accompanist to be less reliant on the performer, playing the accompaniment as expected until they had received more significant evidence that the accompanist had deviated from the score.

The initial findings in this work form the basis for a fruitful avenue for further investigation that could make significant contributions to the interdisciplinary area of computer/human musical interaction. This area has been neglected thus far in score following research but that must be addressed if the accompaniment produced is to be judged musically acceptable by human musicians. The work on *entrainment* (synchronisation) by Clayton et al (2005) looks to be highly relevant here and worthy of further investigation in this context, as does the discussion of interaction between jazz performers by Schögler (2003).

Concluding Remarks

This research has examined the effectiveness of Hidden Markov Models for score following

and concluded that they are a useful tool with which to implement score following systems.

During the lifetime of this work, HMM-based artificial accompanists have been developed in the interactive real-time music processing environment Max/MSP. These artificial accompanists incorporate various enhancements such as beat tracking, the handling of longer scores and the ability to produce complex accompaniment that changes whilst the soloist remains in a particular state.

During development, a Hidden Markov Model structure was partially implemented in Max/MSP to model the scores and to carry out the Viterbi algorithm.

These artificial accompanists are able to determine which HMM state the soloist is currently in, by analysing what the soloist has just played against a specified score. They can then play the appropriate accompaniment for that state.

Performances by each artificial accompanist have been evaluated subjectively by testers of varying musical ability and experience, and also by the objective criteria that was used to evaluate artificial accompanists at the Music Information Retrieval Evaluation eXchange conference of 2006.

Overall the artificial accompanists have been able to produce real-time accompaniment to a human soloist, playing one of three different pieces, of varying complexity. In most cases the accompaniment was musically appropriate throughout the performance of the piece, even when the soloist performer deviated from the score by making errors or adding embellishments to the music performed. An important avenue for further work has been highlighted as a result of this work: the need to reflect on the interaction between soloist and accompanist in performance, for optimum musicality.

Acknowledgments. We would like to thank Kinnell Anderson, Peter Nelson, David Murray-Rust and Michael Edwards at the University of Edinburgh, and Christopher Raphael at Indiana University, for their advice and practical help. The time and effort given by those involved in the testing phase of this work are also much appreciated.

References

- Cano, P., Loscos, A. and Bonada, J. (1999) *Score Performance Matching using HMMs*. Proceedings of the 1999 ICMC, China.
- Clayton, M, Sager, R. and Will, U. (2005) *In time with the music: the concept of entrainment and its significance for ethnomusicology*. European Meetings in Ethnomusicology 11 (ESEM Counterpoint 1)
- Cont, A. and Schwarz, D. (2006) *Score Following Proposal*. http://www.music-ir.org/mirex2006/index.php/Score_Following_Proposal. Accessed August 2007.
- Dannenberg, R. B. (1984) *An On-line Algorithm for Real-time Accompaniment*. Proceedings of the 1984 ICMC, France.
- Dannenberg, R. B. (2000) *Artificial Intelligence, Machine Learning, and Music Understanding*. Proceedings of the 2000 Brazilian Symposium on Computer Music: Arquivos do Simpósio Brasileiro de Computação Musical (SBCM).
- Durbin, R., Eddy, S., Krogh, A. and Mitchison, G. (1998) *Biological Sequence Analysis*. Cambridge, UK: Cambridge University Press.
- Gouyon, F. and Dixon, S. (2005) *A Review of Automatic Rhythm Description Systems*. Computer Music Journal, 29(1): 34-54.
- Grubb, L. and Dannenberg, R. B. (1998) *Enhanced Vocal Performance Tracking Using Multiple Information Sources*. Proceedings of the 1998 ICMC, USA.
- Jordanous, A. (2007) *Score Following: An Artificially Intelligent Musical Accompanist*. Master's thesis, University of Edinburgh. <http://www.inf.ed.ac.uk/publications/thesis/online/IM070498.pdf>. Accessed April 2008.
- Orio, N. and Dechelle, F. (2001) *Score Following Using Spectral Analysis and Hidden Markov Models*. Proceedings of the 2001 ICMC, Cuba.
- Orio, N., Lemouton, S., Schwarz, D. and Schnell, N. (2003) *Score Following: State of the Art and New Developments*. Proceedings of the 2003 NIME, Canada.
- Pardo, B. and Birmingham, W. (2005) *Modeling Form for On-line Following of Musical Performances*. Proceedings of the Twentieth National Conference on Artificial Intelligence, Pittsburgh, Pennsylvania.
- Rabiner, L. R. (1989) *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. Proceedings of the IEEE, vol. 77(2): 257-286.
- Raphael, C. (1999) *Automatic Segmentation of Acoustic Musical Signals Using Hidden Markov Models*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21(4):pp. 360-370.

- Raphael, C. (2001) *Music Plus One: A System for Flexible and Expressive Musical Accompaniment*. Proceedings of the 2001 ICMC, Cuba.
- Schögler, B. W. (2003) *The pulse of Communication in Improvised Music*. Proceedings of ESCOM conference, Germany
- Schwarz, D., Orio, N. and Schnell, N. (2004) *Robust Polyphonic Midi Score Following with Hidden Markov Models*. Proceedings of the 2004 ICMC, USA.
- Vercoe, B. L. (1984) *The Synthetic Performer in the Context of Live Performance*. Proceedings of the 1984 ICMC, France.

¹ The definition of this in Cont and Schwarz (2006) is slightly confusing: "Difference between detection time and the time the system sees the audio" but our interpretation of the latency measure is as described in the main text

² This rating system rated from 0/5 (the accompaniment played bears no resemblance whatsoever to what should have been played) to 5/5 (flawless accompaniment, indistinguishable from or better than the accompaniment that an expert human accompanist would play)

³ HMM-based note/signal artificial accompanist, described at http://www.music-ir.org/evaluation/MIREX/2006_abstracts/SF_cont.pdf

⁴ Dynamic programming-based note artificial accompanist based on Dannenberg (1984), described at http://www.music-ir.org/evaluation/MIREX/2006_abstracts/SF_puckette.pdf

⁵ Details of how these extensions could be implemented can be found in Jordanous (2007).