Journal of New Music Research

ISSN: 0929-8215 (Print) 1744-5027 (Online) Journal homepage: http://www.tandfonline.com/loi/nnmr20

Investigating the Role of Score Following in Automatic Musical Accompaniment

Anna Jordanous & Alan Smaill

To cite this article: Anna Jordanous & Alan Smaill (2009) Investigating the Role of Score Following in Automatic Musical Accompaniment, Journal of New Music Research, 38:2, 197-209, DOI: 10.1080/09298210903180245

To link to this article: http://dx.doi.org/10.1080/09298210903180245

Journal of New Music Research
2009, Vol. 38, No. 2, pp. 197–209

Routledge
Taylor & Francis Group

Investigating the Role of Score Following in Automatic Musical Accompaniment

Anna Jordanous[1] and Alan Smaill[2]

[1]University of Sussex, UK; [2]University of Edinburgh, UK

Abstract

When suitable accompanists are not available to a soloist musician, an alternative possibility is to use computer-generated accompaniment. A computer accompanist should interact with the soloist and adapt to the soloist's playing as a human accompanist would, both reacting to expressive nuances of tempo and to unintentional errors such as wrong or mistimed notes. Over the past 25 years, accompaniment systems have been developed, all of which employ some form of score following: the process of following a musician's progress through the score of a piece during performance. This work considers the role of score following in automatic accompaniment. In this investigation we developed a computer accompanist that employs score following. Our computer musician uses Hidden Markov Models to model the score by metrical structure and to provide accompaniment to a soloist playing monophonic music in real time, as the soloist is playing. Working with MIDI input/output, it tracks tempo fluctuations, anticipates the soloist's next note and supports some amount of unintentional deviation from the score. Qualitative evaluation, by human testers, and quantitative evaluation, using measurable criteria taken from MIREX, reported that the system performs adequately. We then used interviews with eight human accompanists to consider how well a score following system models the accompaniment process. This evaluation raises questions about the musical interaction between soloist and accompanist that have received relatively little attention. The information we gathered from interviews suggests the importance of other aspects of accompaniment, such as the sharing of shape of the performance between musicians, rather than treating the accompanist as purely subservient. We discuss the implications of these issues for the design of automated accompanists.

1. Introduction

Accompanists may not always be available when needed, or available accompanists may not have sufficient technical ability to provide adequate accompaniment. A solution for many musicians is to make use of recorded or computer-generated accompaniment where the accompaniment is static, i.e. never changing from one performance to another. This forces the musician to adapt their playing to synchronize with the accompaniment. It is more natural for the musician, though, if the accompaniment adapts to the performer, particularly as a musician's playing tends to deviate from a 'deadpan' mechanical performance (for stylistic or unintentional reasons). Raphael (2001) describes this change in adaptation as moving from 'music minus one' to 'music plus one'.

To dynamically synchronize the accompaniment with the performance by the musician, the accompanist should track the performer's progress through the score of the piece as they play. *Score following* is the process whereby 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. The purpose of the research outlined in this paper is to investigate the role of score following in automatic accompaniment.

Correspondence: Anna Jordanous, Music Informatics Research Centre/Creative Systems Lab, School of Informatics, University of Sussex, Brighton, BN1 9QJ, UK. E-mail: a.k.jordanous@sussex.ac.uk

DOI: 10.1080/09298210903180245 © 2009 Taylor & Francis

In live performance, score following must be *on-line*: in other words it has to happen in real time, producing accompaniment in time with the soloist's playing. This places extra challenges for the score follower, which has a more limited amount of information available for analysis: only the notes that have been played so far, as opposed to having the whole performance to analyse. Dannenberg and Raphael (2006) discuss this further.

1.1 Previous score following research

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 (1997, 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 (Cano et al., 1999; Orio & Dechelle, 2001; Raphael, 1999, 2001; Orio et al., 2003; Schwarz et al., 2004; Pardo & Birmingham, 2005, Cont & Schwarz 2006b, Macrae & Dixon 2008). Other methods have also been attempted in recent years, such as using belief networks (Raphael, 2000) or graphical models (Raphael, 2004), but Hidden Markov Models are most widely adopted to date.

1.2 Details of the work presented in this paper

During the course of this research, a score follower accompaniment system was developed as a practical tool to assist our investigation. We felt that it was misguided to examine score following without having had experience devising score following systems of our own.

We tested this system qualitatively and quantitatively, then considered the system in the context of how human musicians perform accompaniment. This comparison was made using data gathered from interviews with eight accompanists, on their accompaniment experience and strategies. These steps led to valuable insights for future research in automatic accompaniment systems.

1.2.1 Implementation of an artificial accompanist

Our score follower system was developed over a period of three months; as such, we make no claims that it is as sophisticated as the current state-of-the-art in score following (which we believe to be demonstrated for example by Christopher Raphael's work and by the work at IRCAM by Arshia Cont, Diemo Schwarz and their colleagues). We wished, however, to examine a novel solution to the problem Raphael (2004) highlights: how to represent note durations in a model of the score. Our beat-based model is detailed below.

As we wished to focus our attention on the use of score following, we chose to deal exclusively with MIDI input/output for our score following system, rather than adding the necessary complexity to our system to be capable of processing audio signals.

Taking a lead from the current state of research in score following, our accompanist system uses a Hidden Markov Model representation of the piece being played, to follow the soloist's progress through the musical score and provide accompaniment in real-time. We used a novel approach in fitting Hidden Markov Models to the music being performed: dividing the music into HMM states that represent minimal temporal units of the soloist's part, as opposed to identifying significant events in the score, as in Orio et al. (2003), or Cano et al.'s (1999) use of HMMs to process the incoming audio signal. In the examples treated, each note length (ignoring trills and other ornaments) is a multiple of some length not shorter than a quaver/eighth note; hence each state represents the appropriate minimal temporal unit, corresponding to at most one note in the soloist's score (although one state may correspond to many notes in the accompanist's part).

The system employs simple *beat tracking* to allow for fluctuations in tempo and to be able to anticipate the soloist's next note.

Using HMM states to represent temporal units meant that tempo measurements could be linked directly to the scored notes via the HMM, making beat tracking simple to implement. This representation also simplified the problem of how to model the length of notes, as the ideal length of each note in the score was explicitly represented by the number of HMM states used to model that note's length as a multiple of the basic unit length. Should a soloist play a note for longer (or shorter) than as written in the score, the HMM will consider the soloist to have passed through some extra state(s) (or skipped some state(s)).

While we concede that this approach usually requires a greater number of HMM states than a note-based approach, we feel this is justified by the gains to be had in implementing beat tracking, direct note length modelling and the ease with which we can implement an HMM of a musical score, having at most one possible observation from the soloist per state as opposed to multiple observations per state. To the best of our knowledge, using HMMs to model scores by beat-fraction rather than by note or musical event has not been attempted before in score following research.

1.2.2 Evaluation of the artificial accompanist

To test our score follower for the purposes of accompaniment, and suggest more general reflections on the role of score following in automatic accompaniment, testers of varying musical ability and experience gave qualitative evaluations of the artificial accompanist's performance. The artificial accompanist was also tested extensively using quantitative criteria developed at the

Music Information Retrieval Evaluation eXchange conferences (Cont et al., 2007).

We then considered in more general terms how the score following accompanist compares to human accompanists, interviewing eight human accompanists and comparing their strategies and experiences with the practices used by the artificial accompanist. This evaluation method has not been used in score following research prior to our work, as far as we are aware. We suggest this is an overlooked omission for any work which aims to consider how best to produce an artificially intelligent accompanist that could (if necessary) replace a human accompanist. The insights we gained from examining human musicians' strategies for accompaniment were highly valuable in finding potentially promising lines of future work in the modelling of human accompaniment. We highlight comments from these interviews, summarize the views expressed, and suggest design possibilities for artificial accompanists based on this experience.

2. Developing an artificial accompanist system

The artificial accompanist system was developed in Max/MSP, a programmable music processing environment. Monophonic MIDI input from the soloist was through a MIDI keyboard interface, with MIDI output being generated by the system as accompaniment. The accompaniment system used our implementation of a Hidden Markov Model.

2.1 Hidden Markov Models

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. The comprehensive tutorial by Rabiner (1989) provides an excellent introduction to HMMs.

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, for example biological sequence analysis (Durbin et al., 1998), speech recognition (Rabiner, 1989), as well as in music.

2.1.1 Score following using Hidden Markov Models

A musical score is divided into a sequence of musical events (for example modelling each note as a musical event). The artificial accompanist uses a pre-defined Hidden Markov Model representation of these events to estimate what state the performer is most likely to have reached, i.e. where in the score the performer currently is. It does this by using an algorithm such as Viterbi, outlined by Rabiner (1989), to analyse musical input from the soloist, finding the most probable state sequence that generated the given sequence of observations (notes played by the soloist).

Figures 1 and 2 show how a score can initially be matched to a Hidden Markov Model. The score in Figure 1, with the soloist part in the top stave and the accompaniment in the lower stave, is subdivided into HMM states: in this example each state represents one beat. Figure 2 shows these states, together with the observations for each state and the state transitions, for a performance that never deviates from what is scored. The model in Figure 2 is then developed further to account for performances that may deviate from the score; this is described below and illustrated in Figures 3 and 4.

2.1.2 Details of the Hidden Markov Model representation used

The work presented here places an emphasis on modelling the musical structure of a piece by using HMM states to represent individual minimal temporal units of the soloist's part, an approach to score following which to the best of the authors' knowledge has not been tried before.

This approach minimizes any reliance on us identifying key events in the score accurately and rigorously, in contrast to the approach where HMM states model important note-related events in the score.[1] Additionally it considerably simplifies the implementation of beat tracking; if one state represents a standard, predetermined temporal unit, then the path taken through the Hidden Markov Model by the soloist can be used for direct tempo measurement. The beat tracking implementation is described in more detail below.

There is another advantage of our use of minimal temporal units to subdivide the score into HMM states. By definition, there can only ever be at most one soloist note

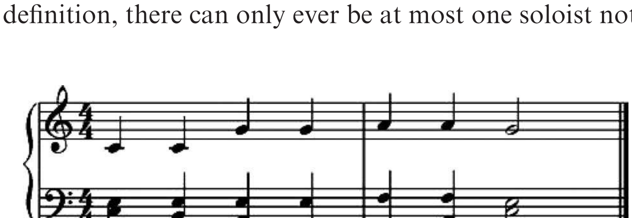

Fig. 1. An example melody (a short extract from a traditional tune: 'Twinkle, Twinkle Little Star', used as Melody 1 for this score follower) with the soloist line in the top stave and the accompaniment in the bottom stave.

[1]Orio et al. (2003) describe score following systems taking a note-based approach.

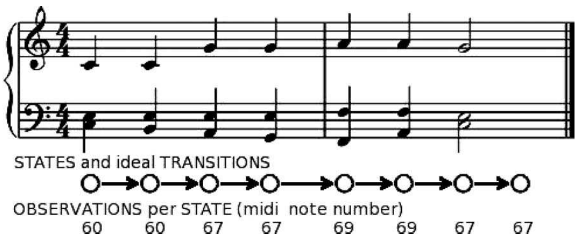

Fig. 2. A simplified version of the Hidden Markov Model for the melody in Figure 1, showing states, transitions and observations for the melody as it is scored (i.e. not yet allowing for any deviations from the score).

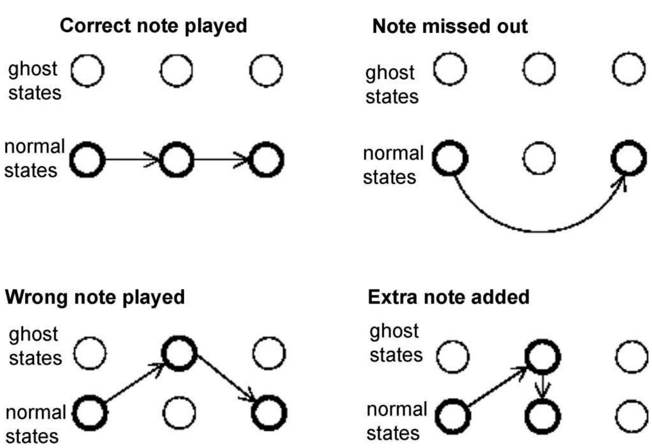

Fig. 3. Typical deviations from a score and the Hidden Markov Model state transitions associated with these deviations (taken from: Orio & Dechelle, 2001).

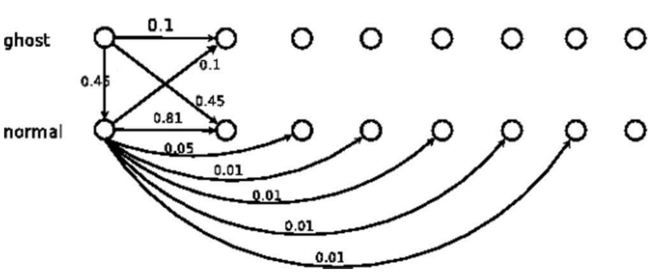

Fig. 4. All allowed transitions from the first normal/ghost state pair, with associated probabilities.

played during one such temporal unit. Therefore, while the soloist is playing their part as scored, there will be at most one observation from the soloist to determine the soloist's route through the Hidden Markov Model. If the soloist does play more than one note during the time-span of this temporal unit, this is either an indication of the soloist changing tempo (which is captured by the beat tracking part of the system, working in conjunction with the HMM) or it is an error or embellishment by the soloist; a clear indication that the soloist has departed from the score (this is dealt with by the use of different types of states, as outlined in the next paragraph). If we had allowed for more than one soloist observation per state when constructing HMM representations of the score, a number of complications would have arisen in how to treat multiple observations per state (most notably: how can the system be sure that a state has been reached, until the soloist has finished playing all the observations that correspond to that state?). Our choice of state representation for the score allows us to bypass these complications in a simple manner.

2.1.3 Normal states and ghost states

Whilst the Hidden Markov 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 states that the soloist passes through if they are playing the piece as written in the score. Ghost states represent states reached by the soloist if they have deviated from the score at that point. Figure 3, adapted from Orio and Dechelle (2001), represents different types of transitions through normal and ghost states, corresponding to different performances by the soloist.

2.1.4 Hidden Markov Model probabilities

The choice of transition probabilities was done by hand, but was guided by general principles, and it should be possible to generate these on the basis of the score information automatically. The general scheme associates a high probability of moving to the next normal state, as expected, but allowing a small probability that a ghost state is entered, and also a smaller probability of moving to future normal states. From each ghost state, transitions are allowed to the corresponding normal state, the next ghost state and also to other states, to allow for repetition/skipping parts out.

The transition probabilities were higher for states near each other in the score and lower for states located further apart in the score. Similarly the initial probabilities were highest for the first normal state, as people are most likely to start with the first note of the piece. It is possible here to allow for other patterns of error, such as ignoring a repeat sign, or turning two pages at once, by adjusting probabilities between the states involved.

Figure 4 demonstrates the general pattern of transition probabilities; details of the exact probabilities used can be found in Jordanous (2007).

The output associated with each hidden state is the accompaniment for the associated point in the score. Transitions between each state model the movement from one point to another; so the transition probabilities were highest for transitions between consecutive normal states (which model the soloist part).

All HMM probabilities were set in advance according to the above heuristics. Currently no training has been implemented for the accompanist to 'learn' the HMM probabilities; however this would be a useful future addition to improve accuracy, particularly for an individual performer or piece.

2.2 Beat tracking: For monitoring the soloist's tempo

A human accompanist would not wait for every note to be played by the soloist before playing accompaniment. Instead they anticipate that the soloist will move onto the next note in the score and play the appropriate accompaniment, then use the incoming information from the soloist to update their belief of where the soloist is in the score and adjust their accompaniment if necessary.

In a similar fashion, this system uses the Hidden Markov Model representation to work out what the next sequential state is, playing the accompaniment for that state at the time it expects the next state to occur. As it receives and processes the soloist's actual input and locates the HMM state that the soloist has actually reached, it adjusts the accompaniment if necessary.

A soloist will naturally incorporate expressive features in the playing, involving shaping of the tempo and intensity of the playing in ways not explicitly represented in the score. As Eric Clarke (2004) remarks, both cultural norms relating to particular styles, and individual aspects of interpretation may be involved here.

The system incorporates a simple version of beat tracking. This allows small tempo fluctuations to be tracked, and the soloist's output to be anticipated in a timely fashion. Modelling the score by temporal units assisted us greatly with including beat tracking in the accompanist. Our implementation was simpler than much recent beat tracking work (Gouyon & Dixon, 2005) but was effective.

The accompanist used an internal tempo measure that was continually adjusted to match the soloist's estimated current tempo, using a local window of notes recently played by the soloist and measuring the time in between those notes (relative to the notes' expected durations). If the soloist is currently judged to be in a ghost state (i.e. they have deviated from the score), then the last input is not considered as valid for use in updating the tempo. If, though, the soloist is currently judged to be in a normal state (i.e. they can be found on the score), then the score follower works out how long the previous note should have been and compares this with the actual length of the last note. The current tempo is based on an average of the recent (valid) tempo observations. The largest and smallest tempo observations are ignored and a mean is taken of the remaining tempo observations, to generate an estimate of the current tempo.

2.3 Controlling dynamics of the performance

The system can track the volume of the soloist's playing using MIDI information and replicate that volume in the dynamic level of the accompaniment output, playing the accompaniment at a very slightly lower volume than the soloist. In this way the system allows the soloist line to be prominent but also matches the dynamic markings of their playing. We felt it was more important to be responsive to the soloist's dynamic interpretations than to allow the accompanist to play at a dynamic marking independent of the soloist's dynamics.

2.4 Performance repertoire

Three melodies were selected for performance by a human soloist and the artificial accompanist. Extracts from Melodies 1, 2 and 3 can be seen in Figures 1, 5 and 6 respectively. Full scores can be found in Jordanous (2007).

- **Melody 1** The first, from the traditional melody 'Twinkle Twinkle Little Star', was the most simple. It had a completely homophonic accompaniment (arranged by the authors), 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. As the minimal metrical unit in the soloist's part was a crotchet/quarter note, each HMM state corresponded to one beat. Two different accompaniments were arranged by the authors for this melody: an accompaniment with no movement independent of the soloist's movement, and a second more complex accompaniment where the accompaniment included part movement during a single HMM state. An example can be seen in bar 1 beat 4 of the extract in Figure 5, where the second of the two quavers (eighth notes) must be timed half way through that HMM state, using the beat tracking information.
- **Melody 3** 'Danse Macabre' (Saint-Saëns) was selected specifically as a more challenging solo melody to track the soloist through, as it incorporates much repetition of note sequences and some stylistic variation in note lengths. The extract in Figure 6 is taken from an arrangement of 'Danse Macabre' for baritone saxophone and piano accompaniment, by Anne Christopherson.

For each piece, a Hidden Markov Model was constructed as described above, with the scored notes in the melody used as the observations connected to transitions between sequential *normal* states. The level of complexity of the melodies and their accompaniments varied across the three melodies, with the more

Fig. 5. An extract from Melody 2: 'All I Ask of You' by Andrew Lloyd-Webber (soloist line in the top stave, accompaniment in the lower stave, arranged by the authors).

Fig. 6. An extract from Melody 3: part of an arrangement from 'Danse Macabre' (Saint-Saëns) for solo baritone saxophone and piano accompaniment, by Anne Christopherson.

challenging melody (Melody 3) including much repetition of note sequences.

3. Evaluating the accompanist system's performance

The overall aim of a competent accompanist should be to provide musical and accurate accompaniment, interacting with the performer in real time.

Our artificial accompanist was evaluated both objectively and subjectively. The system was judged against measurable criteria originally constructed in 2006 by score following experts to test the latest research efforts (Cont & Schwarz, 2006a; Cont et al., 2007). As well as this testing, the artificial accompanist was tested by musicians of varying musical ability and experience, who gave their opinions on the quality of accompaniment provided.

Several parameters of the artificial accompanist were explored during testing, for example either including beat tracking or with no beat tracking (instead asking the tester to play to a metronome speed), or changing the order of the HMM, i.e. the number of historical observations used to track the soloist.

The accompanist was also evaluated according to how closely it matched the approaches and strategies used by a range of human accompanists. This is described in Section 4.

3.1 Methodology for quantitative evaluation

Taking testing criteria from the 2006 Music Information Retrieval Evaluation eXchange (MIREX) conference (Cont et al., 2007), our quantitative evaluation measured the precision and timeliness of the system in tracking the performer, by measuring for each piece:

- Event Count: the number of musical events included in the played melody (i.e. the number of musical events for which the score follower has to estimate a state);
- Number of Notes Missed (and Missed Note %): scored notes that the score follower does not recognize at all, or which are recognized but with an offset of greater than 2000 ms;
- Number of False Positive identifications (and False Positive %): scored notes that the score follower only recognizes after a delay of greater than 2000 ms (also included in Number of Notes Missed);
- Mean and Standard Deviation Offset: the time between an input note occurring and its detection by the system;
- Average Latency: the time between an input note being detected by the system and the accompaniment note onset.

Additionally there are two overall summary measures:

- Total precision: the percentage of correctly detected notes overall (i.e. all score followers' results added together);
- Piecewise precision: the mean of the percentage of correctly detected score notes by each score follower.

Five tests for each melody were carried out, incorporating varying degrees of deviance from the score in performance.

3.2 Methodology for qualitative evaluation

In addition to testing the artificial accompanist against objective measurable criteria, the accompanist was evaluated by four human musicians of different levels of musical competence and experience. The testers were presented with five versions of the artificial accompanist, in order of increasing complexity.

In each test, the testers were asked first to play the melody as correctly as they could, then to include some deviations from the scored melody. 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 comment 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.

3.3 Results and discussion of evaluation

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

3.3.1 Comparisons to other score following systems

As we used the methodology developed for evaluation of a Score Following task at MIREX, it is interesting to see how our results compare to benchmarks set by systems previously presented at MIREX, although as different repertoire and input methods were used at MIREX, we must stress that these can only be very general comparisons. We use the MIREX evaluation criteria (Cont et al., 2007) but not the same evaluation data, though in future work it would be useful to see how our score follower performs on the MIREX evaluation data, for a more robust comparison. Table 1 shows that the artificial accompanist developed in this research performed favourably overall, relative to the two artificial accompanists analysed at MIREX 2006 and 2008[2] (Cont & Schwarz, 2006b; Puckette, 2006; Montecchio & Orio, 2008; Macrae & Dixon, 2008). It was well outperformed by the leading system in these tests, the IRCAM accompanist presented at MIREX 2006 by Cont and

Table 1. Summary of results of score followers tested at MIREX 2006–2008, to see the levels of performance currently being achieved by artificial accompanists. NB: We include our system in this table for some general comparison, but stress that our system was tested on different repertoire.

Authors	Total precision	Piecewise precision
Cont & Schwarz (2006b)	82.90%	90.06%
Puckette (2006)	29.75%	69.74%
Montecchio & Orio (2008): system 1	n/a	66.50%
Montecchio & Orio (2008): system 2	n/a	55.26%
Macrae (2008): system 1	n/a	21.22%
Macrae (2008): system 2	n/a	22.85%
This work	60.89%	54.04%

[2]MIREX did not host a Score Following task in 2007.

Schwarz, however the development time of our work was considerably shorter than for IRCAM's longer term project.

During comparison, it was interesting to see the variance in the accuracy of the MIREX artificial accompanists, depending on what piece is being played. This was also observed over different pieces for our artificial accompanist.

3.3.2 Overall performance of the beat-based score follower

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

Our artificial accompanist in general performed better with musicians of lower rather than higher ability, responding better to inconsistent tempos and errors, as opposed to decorative embellishments. This is probably partly due to a slight bias in how we set the HMM probabilities, towards recovering from errors rather than dealing with decorations and embellishments. In most tests the artificial accompanist performed well in responding to different types of tester errors and coped with note embellishments to a certain degree.

We have chosen to link in the state with minimal temporal units. The shortest note duration in Melody 1 is a crotchet/quarter note. In Melodies 2 and 3 the shortest note duration is a quaver/eighth note. We are aware that scores including notes with shorter note durations or polyrhythms would require a large increase in the number of states necessary to represent the score, so acknowledge that these melodies are not fully representative of the space of possible scores and that this is an area that would require further work.

3.3.3 Latency and accuracy issues

Unsurprisingly, when the artificial accompanist incorporated beat tracking, there were higher latency measurements for receiving and processing the soloist's playing (a difference of approximately 200 ms in general). This is due to the extra processing involved.

Including beat tracking in the accompanist led to the testers judging the accompanist to be more accurate and musical, as it adjusted to their playing well. This inclusion also led to more accuracy being reported in the quantitative measurements, probably because the system relied less on the soloist playing in a strict metronomic manner and could adjust to occasions where the soloist did not time their playing exactly according to the given tempo.

Using a larger number of observations for the Viterbi algorithm (four observations instead of three) showed much improvement in performance accuracy. This shows the improvements possible when using a greater amount of information from the soloist. Latency measurements associated with the more detailed calculations were, however, considerably higher. This finding was reflected in user testing, where testers consistently reported that this version of the accompanist lagged behind the soloist.

The overall accuracy measurements for some tests on the most complex 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 did locate the performer correctly in the score, but not quickly enough to produce musically accurate accompaniment.

3.3.4 Detecting tempo changes

In evaluation, the testers generally judged the artificial accompanist as being able to detect changes in tempo rapidly, although this could be improved further. This functionality of the artificial accompanist is related to the ability to track the performer accurately through the piece, so if the artificial accompanist tracks the soloist through the score more competently, there is an associated improvement in tempo tracking.

3.4 Reflections on the role of the accompanist

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. All four testers on occasion would attempt to synchronize with the accompanist when the accompaniment was not quite correct. Some tried to assist the accompanist in these situations by giving musical cues such as accenting the first beat of a bar. One tester in particular stressed in their feedback how ensemble performance is a co-operative process; the human will react instinctively to the computer's playing.

These comments led us to consider an extra stage of evaluation: how do the strategies and approaches used by our artificial accompanist compare to those of a human accompanist? This is discussed in the next section.

4. Comparing the artificial accompanist to human accompanists

A common finding during testing was that the artificial accompanist was too reliant on following the performer, rather than using knowledge of where it had previously been in the score to play a more continuous form of accompaniment. Testers also noted that the accompanist did not respond to cues from the soloist.

Roger Dannenberg has commented on a similar finding in an ensemble situation (Dannenberg, 2000, p. 3):

> 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.

This feedback raises the point: to what extent can Score Following successfully be used to simulate human musicians performing accompaniment? In improving our accompanist should our efforts concentrate on aspects other than exploring different uses of Hidden Markov Model or alternative Score Following implementations? To consider this, we investigated how closely the artificial accompanist represented accompaniment practice by human musicians. Throughout this section, all anonymous quotes are from these interviews.

4.1 Methodology for comparison to human accompaniment

We interviewed eight musical accompanists; representing a range of accompaniment experience. Musicians were chosen who were familiar with the accompaniment scenario considered in this work, i.e. a soloist playing to accompaniment, where the accompanist is given a score to play from that contains both the accompanist part and the soloist part.

The eight musicians had between 2–45 years accompaniment experience (mean 14.8 years, standard deviation 13.5) and between 11–58 years experience playing their accompaniment instrument (mean 23.1 years, standard deviation 14.7). All eight accompanists had performed accompaniment on piano and/or keyboard; three also had experience accompanying on organ. Five of these musicians had additional conducting experience, directing a group of musicians to accompany solo performances.

Typical accompaniment scenario for these accompanists varied widely from professional musicians to novices, in a variety of genres from Baroque through to popular music. Academic qualifications of the accompanists varied, from no official qualifications to a postgraduate degree in accompaniment from a leading British music college.

Each accompanist was asked about their approach to accompaniment and the strategies they used. The primary areas discussed were:

- Reflections on their personal approach to accompaniment;
- How they synchronize their playing with their soloist;
- To what extent they would be aware of deviations from what is written in the score;
- How they would deal with such deviations.

General conclusions arising from these interviews could then be used for comparison to our artificial accompanist.

4.2 Similarities between the artificial system and human musicians

We found a number of similarities between the musical strategies described by the accompanists interviewed and the heuristics underpinning the computer accompanist. Table 2 summarizes the similarities highlighted in the interviews.

4.2.1 Score following using the soloist's playing

All eight accompanists interviewed, without exception, commented frequently on how important it was to keep track of where the soloist had reached in their performance. Moreover, repeatedly they stressed how important it was to listen to the soloist in the context of what had previously been played; replicated in our computer accompanist with the use of Hidden Markov Models, which capture the contextual nature of the sequence of notes coming from the soloist.

> If you're not listening to them, if you're not following them, if you're focusing on playing what you're supposed to be playing, it doesn't happen.

> Listen think listen think listen think and then repeat until you are sure you know where they are.

> I think it's still accepted that, as the pianist has got all of the parts in front of his eyes, he can see what's happening.

As such a high importance was placed on following a soloist's playing in order to accompany them most appropriately, this acts as strong evidence that incorporating

Score Following in artificially intelligent accompaniment is crucial for musical success.

4.2.2 Deviations from the score by soloists

A fundamental issue surrounding this work is that soloists very rarely, if ever, play a piece exactly as scored. This was echoed in the interviews. The general consensus was that more advanced musicians tended to make intentional stylistic deviations from the score whilst less gifted musicians were more likely to make unintentional mistakes. Typical deviations included adding embellishments and ornamentations, or deliberately varying their interpretation of rhythms, tempo, musical feel and expression, as well as playing wrong notes, missing out part of the music and uncontrolled variation in tempo (particularly speeding up when nervous).

4.2.3 Problems arising from practising with static accompaniment

One of the original motivations for this project was to provide an alternative to practising with a fixed, static accompaniment for musicians who have no suitable accompanist available. Two interviewees specifically mentioned comments on problems they had encountered when working with performers who had previously practised with pre-recorded accompaniment. Both described the difficulties that these performers had in adjusting to live accompanists after such practice. Though it is not uncommon to face problems adjusting to a new accompanist, in one situation the performers had developed timing issues as a result of adapting their performance to the static recording.

Table 2. Similarities between human accompanist strategies and this artificial accompanist, as highlighted in interviews with eight human accompanists.

Common strategy	Accompanists mentioning this strategy	Total no. of comments
Score Following using the soloist's playing (especially when out of sync)	8 (00, 01, 02, 03, 04, 05, 06, 07)	27
Deviations from the score	8 (00, 01, 02, 03, 04, 05, 06, 07)	31
Problems of practising with fixed accompaniment	2 (01, 06)	3

4.3 Aspects missing from the artificial accompanist

Whilst fundamental similarities exist between the artificial accompanist and humans' accompaniment strategies, many observations raised during the interviews indicate a number of ways in which the artificial accompanist differs from human approaches. A summary is presented in Table 3.

4.3.1 Co-operation between soloist and accompanist

From playing the soloist's part with them or calling out directions to the soloist, to constructing a musical dialogue between soloist and accompanist, the co-operative aspect of performance accompaniment was stressed by all those interviewed. The simple scenario of our accompanist following the soloist, with no musical input or directions of its own, does not reflect the accompanists' descriptions of their own work.

> Everyone makes mistakes, so you pick them up if you're in a position where you can do that.

> There are two people adapting in this situation.

Many mentioned the role of the accompanist in guiding their soloist through the piece, perhaps taking a pedagogical role; some pointed out their input in mutual decisions on stylistic interpretation.

One accompanist questions the prominence of Score Following in accompaniment, whether human or artificial:

> In terms of the ensemble, playing together, it's nice to take a few risks, and you can do that with somebody that you've worked with for a while, and get away from that following kind of idea which you sometimes seem to cultivate as an accompanist, but to accompany means 'to go with', not to follow behind.

Table 3. Differences in accompanist strategies between human accompanists and this artificial accompanist, as highlighted in interviews with eight human accompanists.

Contrasting strategy	Accompanists mentioning this strategy	Total no. of comments
Co-operation between soloist and accompanist	8 (00, 01, 02, 03, 04, 05, 06, 07)	38
Adaptation to different performance scenarios	8 (00, 01, 02, 03, 04, 05, 06, 07)	11
Musical awareness	6 (00, 01, 02, 03, 05, 06, 07)	11
Using score formats other than 'soloist line + accompaniment'	6 (00, 02, 03, 04, 05, 06)	7
Coping with unexpected problems\technical hitches	3 (00, 01, 06)	4

4.3.2 Adapting to different performance scenarios

The approach of having one standard accompaniment process for all scenarios seems flawed, given our findings. Every interview saw discussion of how accompaniment strategies change for different performers. In general this was linked to the performers' ability. This links back to an observation made during qualitative testing: when testers perceived the system to be doing badly in accompanying them, they often changed the way they were playing to give performance cues such as accenting the first beat of the bar, to help the system align back with their playing. Testers continuously did this during their entire testing session, even though they soon realized that this did not trigger any improvement in accompaniment performance.

> When you play with a weak soloist you may have to play more rigid tempos and emphasize key notes and phrases whereas with a more accomplished soloist you need to really understand how they perform the piece and help them to deliver the music the way they see it.

> [With an advanced soloist] I'd concentrate more on putting my musicality across, rather than worrying so much about what she was doing.

4.3.3 Musical awareness

Several of those interviewed described how they use their musical knowledge for heightened awareness during performance; this is beyond the present capabilities of the computer accompanist.

> You're probably more focused on checking that you're in synch at the end of a phrase, it's harder midway through.

> Do they need to slow this down subtly so they can get their fingers round it?

> Try to extract certain patterns, or you aim for the cadence.

4.3.4 The role of the score in musical performance

Our choice of representation assumes that the accompanist is given both a fully written accompaniment and a copy of the soloist line. This is, however not always the case. From figured bass to jazz chord sheets, most accompanists had encountered alternative score formats. A common reaction to not having the soloist's line, however, was:

> I get very disconcerted, actually, if I'm handed an accompaniment that only has my part on it.

4.3.5 Coping with unexpected/technical problems

Two of the accompanists described situations where they had to deal with unforeseen technical problems to do with their equipment malfunctioning or their making a mistake in setting the equipment up. A third interviewee described a situation where the performers had to deal with a fire alarm sounding mid-performance. Coping with unexpected problems is clearly relevant for a system that relies upon technical equipment functioning as it is designed to; robustness of the system is an important issue.[3]

4.4 The artificial accompanist as a model of human musicians

At a basic level, the artificial accompanist does resemble human musicianship to some extent; though testing shows a slight bias towards accompanying novice rather than advanced performers.

Many of the differences we found between computer and human musician are aspects which could feasibly be added to our model. For example, incorporating musicological analysis into the accompanist would increase its musical awareness, or physical aspects of playing a sequence of notes on a given instrument could be taken into account.

Another suggestion which may address some of the above discrepancies is to have an explicit model of the soloist, adaptable for different soloists. This internal model, similar to that used in intelligent tutoring systems, would be able for example to recognize an erratic or unconfident soloist and react appropriately. So the accompanist would have different strategies that could be deployed, either under the soloist's direct control or automatically by building a user model of the soloist and using that to choose the appropriate strategies. This model would allow for different accompaniment strategies to come into play as the soloist becomes more familiar with the music.

There are some musical gestures which are employed not to shape the music, but to enable closer interaction between the performers; for example, an emphasis on the downbeat intended to give the accompanist a strong hint of the position of the start of the bar. Such hints could be recognized for what they are by the artificial accompanist.

It is also clear that in human accompaniment there is a lot of interaction that happens by extra-musical means, where physical gestures help with co-ordination, indicate that the player wants the music to be slower, or louder, and so on. Incorporating information about physical posture thus could permit these sort of cues to be used also. Recent work on *speckled computing*[4] shows that such information can be gathered without an elaborate visual motion capture system, and integrated with conventional computational systems.

---

[3]We make no claims here for how an artificial accompanist should best react in the event of a fire alarm.

[4]Speckled computing project homepage: http://www.specknet.org/, retrieved April 2009.

Using human examples as a guide to implementing artificially intelligent accompaniment looks to be a fruitful way of developing the accompanist further.

5. Conclusions

An artificial accompanist has been developed which can follow a soloist through a score (even if the soloist's performance is occasionally inaccurate or embellished) and play appropriate accompaniment. It does this by matching the soloist's playing to a Hidden Markov Model of that score.

Using HMMs considerably simplified our implementation of score following, by providing a framework in which the artificial accompanist could process the sequence of notes played by the soloist. In particular, the ease with which we could implement an HMM representation and incorporate beat tracking in the accompanist proved the worth of our decision to model the score using temporal units, rather than following the note-based representations described in previous work (such as Orio et al., 2003; Cano et al., 1999).

Performances by the artificial accompanist were 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 conferences of 2006 and 2008. Overall the artificial accompanist was able to produce adequate online accompaniment to a human soloist over a repertoire of three pieces, of varying complexity. For simpler pieces, the accompaniment was generally deemed musically appropriate, even when the soloist deviated from the score by making errors or adding embellishments to the music performed. For more complex pieces, though, latency issues severely disrupted the flow of the artificial accompaniment. 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 re-implementing the system using different software). Also as the state representations depend upon what the shortest note duration in the score is, we acknowledge that there are unresolved issues with our choice of representation, that require addressing in future work: when the score contains notes of very short duration and/or polyrhythms. While our choice of representation is practicable for rhythmically straightforward styles, it could be problematic where the rhythmic combinations are more complex.

The artificial accompanist mirrors human accompanists' strategies in some ways, for example demonstrating the awareness necessary to track the performer through the piece, in the context of what has been heard before. The human accompanists, however, make far greater use of co-operation and musical adaptability, compared to the computer accompanist. This suggests a fruitful direction of further work for improving the success of the artificial musician in accompaniment: that of incorporating musical knowledge in the accompanist strategies and allowing the accompanist more autonomy over what it plays, as opposed to purely following the soloist. Building an internal model of the soloist could be useful to guide the accompanist in its choices of accompaniment strategy, adapting to the individual soloist.

We conclude that a score following system using a beat-based Hidden Markov Model performance can perform reasonably well as an artificial accompanist, although as the accompaniment repertoire becomes more complex, latency issues become more prominent. Furthermore, for increased success in artificially intelligent accompaniment, one should investigate the interactive nature of soloist-accompanist performance, in addition to using score following.

Acknowledgements

We thank Kinnell Anderson, Peter Nelson, Dave Murray-Rust, Nick Collins and Christopher Raphael for their advice and practical help. The comments from two anonymous reviewers, from Emilios Cambouropoulos and from the participants at the Conference for Interdisciplinary Musicology were also very useful.

Those who tested the accompaniment system gave valuable feedback. Similarly the knowledge and experience of the accompanists whom we interviewed played an instrumental part in placing this work in a wider context.

References

Cano, P., Loscos, A., & Bonada, J. (1999). Score performance matching using HMMs. In *Proceedings of the International Computer Music Conference*, China.

Clarke, E. (2004). Empirical methods in the study of performance. In E. Clarke & N. Cook (Eds.), *Empirical Musicology* (pp. 77–102). New York: Oxford University Press.

Cont, A., & Schwarz, D. (2006a). Score following proposal. *Music Information Retrieval Evaluation eXchange Conference*, Canada. Retrieved April 2009, from http://www.music-ir.org/mirex2006/index.php/Score_Following_Proposal

Cont, A., & Schwarz, D. (2006b). Score following at IRCAM. In *Music Information Retrieval Evaluation eXchange Conference*, Canada.

Cont, A., Schwarz, D., Schnell, N., & Raphael, C. (2007). Evaluation of real-time audio-to-score alignment. In *Proceedings of The International Society for Music Information Retrieval Conference*, Austria.

Dannenberg, R.B. (1984). An on-line algorithm for real-time accompaniment. In *Proceedings of the International Computer Music Conference*, France.

Dannenberg, R.B. (2000). Artificial intelligence, machine learning, and music understanding. In *Proceedings of the 2000 Brazilian Symposium on Computer Music: Arquivos do Simposio Brasileiro de Computao Musical*, Brazil.

Dannenberg, R.B., & Raphael, C (2006). Music score alignment and computer accompaniment. In *Communications of the ACM*, *49*(8), 38–43.

Durbin, R., Eddy, S., Krogh, A., & Mitchison, G. (1998). *Biological Sequence Analysis*. Cambridge, UK: Cambridge University Press.

Gouyon, F., & Dixon, S. (2005). A review of automatic rhythm description systems. *Computer Music Journal*, *29*(1), 34–54.

Grubb, L., & Dannenberg, R.B. (1997). A stochastic method of tracking a vocal performer. In *Proceedings of the International Computer Music Conference*, Greece.

Grubb, L., & Dannenberg, R.B. (1998). Enhanced vocal performance tracking using multiple information sources. In *Proceedings of the International Computer Music Conference*, USA.

Jordanous, A. (2007). *Score following: An artificially intelligent musical accompanist* (MSc thesis). Department of Informatics, University of Edinburgh, UK. Retrieved April 2009, from http://www.inf.ed.ac.uk/publications/thesis/online/IM070498.pdf

Macrae, R., & Dixon, S. (2008). Polyphonic score following using on-line time warping. In *Music Information Retrieval Evaluation eXchange Conference*, USA.

Montecchio, N., & Orio, N. (2008). Automatic alignment of music performances with scores aimed at educational applications. *Automated Solutions for Cross Media Content and Multi-channel Distribution*, Florence, Italy.[5]

Orio, N., & Dechelle, F. (2001). Score following using spectral analysis and Hidden Markov Models. In *Proceedings of the International Computer Music Conference*, Cuba.

Orio, N., Lemouton, S., Schwarz, D., & Schnell, N. (2003). Score following: State of the art and new developments. In *Proceedings of New Interfaces for Musical Expression*, Canada.

Pardo, B., & Birmingham, W. (2005). Modeling form for on-line following of musical performances. In *Proceedings of the Twentieth National Conference on Artificial Intelligence*, USA.

Puckette, M. (2006). (Untitled submission for score following task). *Music Information Retrieval Evaluation exchange Conference*, Canada. Retrieved April 2009, from http://www.music-ir.org/evaluation/MIREX/2006_abstracts/SF_puckette.pdf

Rabiner, L.R. (1989). A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286.

Raphael, C. (1999). Automatic segmentation of acoustic musical signals using Hidden Markov Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *21*(4), 360–370.

Raphael, C. (2000). Synthesizing musical accompaniments with Bayesian belief networks. *Journal of New Music Research*, *30*, 59–67.

Raphael, C. (2001). Music plus one: A system for flexible and expressive musical accompaniment. In *Proceedings of the International Computer Music Conference*, Cuba.

Raphael, C. (2004). A hybrid graphical model for aligning polyphonic audio with musical scores. In *Proceedings of The International Society for Music Information Retrieval Conference*, Spain.

Rowe, R. (2004). *Machine Musicianship*. Cambridge, MA: MIT Press.

Schwarz, D., Orio, N., & Schnell, N. (2004). Robust polyphonic midi score following with Hidden Markov Models. In *Proceedings of the International Computer Music Conference*, USA.

Vercoe, B.L. (1984). The synthetic performer in the context of live performance. In *Proceedings of the International Computer Music Conference*, France.

[5]The MIREX abstract for Montecchio and Orio's system was unavailable on the MIREX website at the time of submitting this paper.