The Interactive Robotic Percussionist - New Developments in Form, Mechanics, Perception and Interaction Design

Gil Weinberg
Georgia Tech
840 McMillan St.
Atlanta, Georgia, 30332
(1) 404-894-8939
gilw@gatech.edu

Scott Driscoll
Georgia Tech
840 McMillan St.
Atlanta, Georgia, 30332
(1) 404-894-8939
scott.driscoll@gatech.edu

ABSTRACT
We present new developments in the improvisational robotic percussionist project, aimed at improving human-robot interaction through design, mechanics, and perceptual modeling. Our robot, named Haile, listens to live human players, analyzes perceptual aspects in their playing in real-time, and uses the product of this analysis to play along in a collaborative and improvisatory manner. It is designed to combine the benefits of computational power in algorithmic music with the expression and visual interactivity of acoustic playing. Haile's new features include an anthropomorphic form, a linear-motor based robotic arm, a novel perceptual modeling implementation, and a number of new interaction schemes. The paper begins with an overview of related work and a presentation of goals and challenges based on Haile's original design. We then describe new developments in physical design, mechanics, perceptual implementation, and interaction design, aimed at improving human-robot interactions with Haile. The paper concludes with a description of a user study, conducted in an effort to evaluate the new functionalities and their effectiveness in facilitating expressive musical human-robot interaction. The results of the study show correlation between human's and Haile's rhythmic perception as well as user satisfaction regarding Haile's perceptual and mechanical abilties. The study also indicates areas for improvement such as the need for better timbre and loudness control and more advance and responsive interaction schemes.

Categories and Subject Descriptors
J.5 [**Arts and Humanities**]: Performing arts

General Terms
Performance, Design, Algorithms, Human Factors.

Keywords
Music, Percussion, Perception, Robotics

HRI'07, March 8-11, 2007, Arlington, Virginia, USA.

1. INTRODUCTION
Most computer supported interactive music systems (Winkler 2001) are hampered by their inanimate nature, which does not provide players and audiences with physical and visual cues that are essential for creating expressive musical interactions. For example, motion size often corresponds to loudness and gesture location often relates to pitch. These cues provide visual feedback and help players anticipate and coordinate their playing. Such cues also create a more engaging experience for the audience by providing a visual and physical connection to the sound. Computer supported interactive music systems are also limited by the electronic reproduction and amplification of sound through speakers, which cannot fully capture the richness of acoustic sound. Our approach for addressing these limitations is to utilize a mechanical apparatus that converts digital musical instructions into physically generated acoustic sound. We believe that musical robots can bring together the unique capabilities of computational power with the expression, richness and visual interactivity of physically generated sound. A musical robot can combine algorithmic analysis and response that extend on human capabilities with rich sound and visual gestures that cannot be reproduced by speakers. Our first effort to create such novel human-machine interaction focused on rhythm and involved a simple percussive robotic arm driven by a number of rudimentary perceptual and interaction schemes (Weinberg et. al 2005). In this paper we describe the next stage of this project, which consists of a two-armed anthropomorphic robot named Haile that utilizes advanced interactive drumming techniques, a new perceptual modeling implementation for rhythmic stability and similarity, and a user study that evaluates human-robot interaction with Haile.

2. RELATED WORK
Current research directions in musical robotics focus mostly on sound production and rarely address perceptual aspects of musicianship, such as listening, analysis, improvisation, or group interaction. Such automated musical devices include both Robotic Musical Instruments – mechanical constructions that can be played by live musicians or triggered by pre-recorded sequences such as in (Singer, et al. 2003), (Jordà 2002) or (Dannenberg, et al. 2005) – and Anthropomorphic Musical Robots – hominoid robots that attempt to reproduce the action of human musicians such as in (Takanishi, et al. 1998), (Sony 2003) and (Toyota 2004). Only a few attempts have been made to develop perceptual robots that are controlled by neural networks or other autonomous methods (Baginsky 2004).

The main goal of our project is to embed machine musicianship (Row 2004) in such mechanical devices. In this field, researchers are developing systems that analyze, perform, and compose music based on theoretical foundations in music theory, cognition, artificial intelligence and human-computer-interaction. One of the earliest directions in this area was the Score Follower, in which computers track live soloists and synchronize MIDI (Dannenberg 1984) (Vercoe 1984), and recently audio (Orio 2003) accompaniment to musical input. A more improvisatory approach is taken by systems such as Cypher (Rowe 1992) Voyager (Lewis 2000) or the Continuator (Pachet 2002). In these systems the software analyzes musical input and uses this analysis to generate algorithmic response by controlling a variety of parameters such as melody, harmony, rhythm, timbre, and orchestration. These systems are not designed to generate acoustic sound and remain in the software domain. Due to the percussive nature of the robotic percussionist project, our work is also informed by research in computational modeling of rhythm perception. Lower level cognitive rhythmic modeling addresses detection of percepts such as note onset, tempo, and beat, using audio sources (Puckette 1998) (Scheirer 1998) (Foote, et al. 2001) or MIDI (Winkler 2001). Detection and analysis of higher-level rhythmic percepts include more subjective concepts such as rhythmic stability, similarity, and tension (Paulus, et al. 2002), (Tanguiane 1993), (Coyle, et al. 1998), and (Desain, et al. 2002). Informed by these models, Haile's perceptual and interactive modules are aimed at extracting musical meaning from real-time live drumming, and responding to the acoustic input based on perceptual analysis.

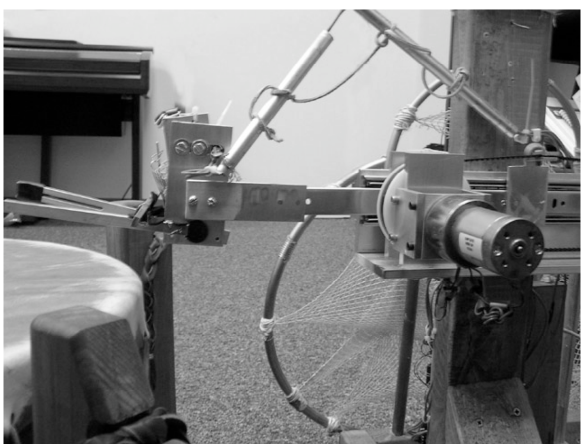

Figure 1. Haile's original one-armed design

3. GOALS AND CHALLANGES

Based on our evaluation of Haile's original prototype (Weinberg et. al 2005) we identified four main areas for further development – the robot's physical form, mechanics, perception, and interaction schemes. Our main challenge in designing Haile's appearance was to provide a more familiar and inviting form in comparison to the original amorphous design (see Figure 1). The robot's shape, construction materials, and the manner in which technology was embedded had to be redesigned to support a more intuitive and engaging musical interaction. In the area of mechanics, our main challenge was to improve the acoustic variety, the dynamic range, and the visual cues provided by the original robotic arm. In perception, we aimed at developing a new algorithmic implementation of rhythmic stability and similarity that could be evaluated in a user study. Based on this perceptual analysis, our goal in interaction design was to develop responsive improvisation algorithms that relate to humans' musical input, using transformative and generative methods both sequentially and synchronously.

4. IMPLEMENTATION

4.1 Physical Form

Haile's original one-arm design did not facilitate the familiar and expressive interactions we aimed for. The design was purely functional and did not communicate the idea that it could interact with humans by listening, analyzing, and reacting. The new form, therefore, had to better convey the robot's interactive purpose as well as to match the aesthetics of the Native American pow-wow drum, a unique multi-player instrument that supports the collaborative nature of the project. An anthropomorphic design made by Clint Cope was chosen to reflect the human-like capabilities of the robot and wood was selected as the primary construction material to match the drum. The new design includes two percussive arms that can strike with varying velocities and move back and forth between the rim and center of the drum. Primarily constructed of plywood cut on a CnC machine, Haile also utilizes metal joints that allow its arms to move and its legs to adjust to different drum heights. While attempting to create an organic look for the robot, it was also important that the technology is not hidden, so that players can see and understand the robot's operation. We, therefore, left the mechanical apparatuses uncovered and embedded a number of LEDs on Haile's body, providing an additional representation of the mechanical actions (see Figure 2).

Figure 2. Haile's new anthropomorphic design

4.2 Mechanics

Haile's original solenoid driven arm was able to generate fast hits but could not produce loud sounds that compare with the amplitude of human's drumming. Moreover, the solenoid's small movements did not provide significantly noticeable visual cues for humans to anticipate and synchronize their gestures. These weaknesses hampered Haile's expression and interaction potential. We, therefore, decided to develop a new arm that would use a powerful linear motor, controlling large movements that are loud and visually noticeable. The new arm is embedded in the new anthropomorphic design as the left arm, and the original solenoid actuated arm as the right arm. Both arms can adjust the sound of strikes in two manners: different pitches are achieved by striking the drumhead in different locations along its radius, and volume is adjusted by hitting with varying velocities. Unlike robotic drumming systems that allow hits at only a few discrete locations (Jordà 2002) (Rae 2005), Haile's arms can strike anywhere on a line between the center and the rim of the drum, moving between the two ends in about 250 ms. To move to different positions over the drumhead, both arms employ a linear slide, belt, gear motor, and potentiometer that provides closed-loop control of the strike location over the drumhead. The right arm's striking mechanism consists of a solenoid driven aluminum stick and a return spring (see Figure 3.), which can strike at a maximum speed of 15 Hz. The more powerful and sophisticated striking mechanism of the left arm uses a linear motor and encoder for closed-loop position and velocity control over the arm height (see Figure 4). The left arm can only strike at about 11Hz, but is able to create much more visible motions, louder volumes, and more controllable variation in volume. By combining the louder and slower left arm with the faster and weaker right arm we aimed at enriching Haile's acoustic richness and diversity.

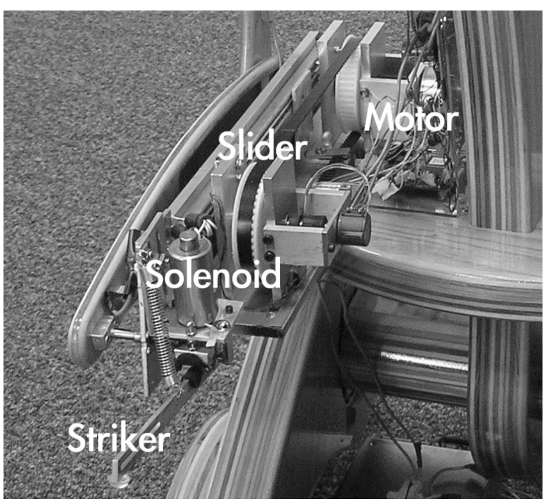

Figure 3. Haile right arm design

In an effort to provide an easy-to-program environment for Haile, we decided to use Max/MSP (Cycling74 2006), an intuitive graphical programming environment that can make the project accessible to composers, performers, and students. The original solenoid based arm communicated with the computer via a USB based Teleo System (MakingThings 2006). Its low-level control was computed within Max/MSP, which required a continuous feed of position updates to the computer. This consumed much of the communication bandwidth as well as processing time on the main computer. The current two-arm mechanism utilizes multiple onboard microprocessors for local low-level control as well as Ethernet communication with the main computer. The new system, therefore, facilitates much faster and more sophisticated control (2ms control loop) and requires only low bandwidth communications with the operating computer. Each arm is locally controlled by an 18F452 PIC microprocessor, both of which receive RS232 communications from a Modtronix Ethernet board (SBC68EC). The Ethernet board receives 3 byte packets from the computer, a control byte and two data bytes. The protocol utilizes an address bit in the control byte to send the information onto the appropriate arm processor. The two data bytes typically contain pitch and velocity set points for each hit, but can also be used to update the control parameters.

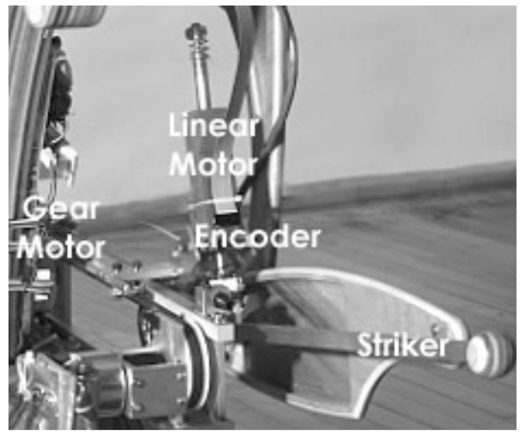

Figure 4: Haile's left arm design.

The onboard PIC microprocessors are responsible for controlling the arms' sliding and hitting mechanisms, ensuring that the impacts occur at the requested velocity and location. In order to allow enough time for the arms to move to the correct location and execute the strokes, a 300ms delay line is implemented between signal reception and impact. It has been shown that rhythmic errors of only 5 ms are detectable by average listeners (Coren, et al, 1984), therefore, it was important to ensure that the delay remained accurate and constant regardless of different hit velocities, allowing us to easily compensate for it in the higher-level interaction application. Both arms store incoming hit commands in a First-In-First-Out queue, moving towards the location of a new note immediately after each hit. Due to its short vertical hitting range, the solenoid driven right arm has a fairly consistent stroke time. We, therefore, implemented a constant delay of 270ms between signal reception and solenoid energization for this arm. The left arm, on the other hand, undergoes much larger movements, which require complex feedback control to ensure that impact occurs exactly 300ms after a hit signal is received. While waiting for incoming notes, the left arm remains about one inch above the surface of the drum. When a new note is received, the arm is raised to a height proportional to the loudness of the hit, and after a delay determined by the desired velocity and elevation, the arm starts descending towards the drumhead under velocity control. After impact, the arm returns back to its standby position above the drumhead. Extremely fast notes utilize a slightly different control mode that makes use of the bounce of the arm in preparation for the next hit. Using this mechanism, the new left arm can control a wide dynamic range in addition to providing performers and viewers with anticipatory visual cues, enhancing expression and enriching the interaction representation.

4.3 Perception

As a test bed for musical human-robot interaction, we developed a number of independent perceptual modules for Haile, which can be embedded in a variety of combinations in compositions, educational activities, and user studies. Haile listens to human drumming via directional microphones installed in hand drums such as djembes, darbukas, or the pow-wow drum itself. Each perceptual module addresses one perceptual aspect, from hit onset, amplitude, and pitch detection, through beat and density analysis, to rhythmic stability and similarity perception. Similarly to the original design, we base our low-level modules for hit onset and amplitude detection on the Max/MSP *bonk~* object (Puckette 1998), and adjust its output to the unique character of the pow-wow drum. In the original design, however, *bonk~*'s frequency band output was insufficient for accurate pitch detection of the pow-wow's low and long reverberating sounds. Since *bonk~* is hard-coded with a 256 point analysis window, the lowest frequency it can analyze is 172Hz – too high for the pow-wow drum, which has a natural frequency of about 60 Hz. Moreover, pitch detection is complicated when high frequency hits are masked by the long decay of the previous low-pitched strikes. To address these issues, we wrote a Max/MSP external object that uses 2048 point FFT frames to determine the magnitude and derivative of lower frequency bins. By taking into account the spectral changes in addition to magnitudes, we can better determine whether energy in a particular frequency band comes from a current hit or from previous ones (see Figure 5).

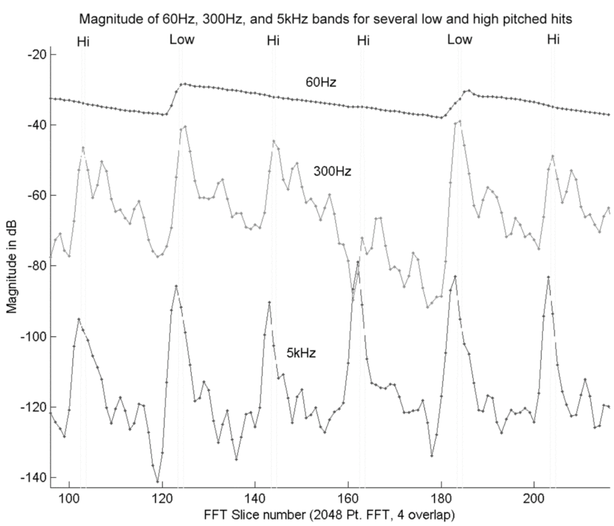

Figure 5: Magnitude plots from a 60Hz, 300Hz, and 5kHz frequency band over several low and high-pitched hits showing the relatively slow decay of the low-pitched hits.

Other relatively low-level perceptual modules added to Haile's perceptual functionalities include beat detection, utilizing Tristan Jehan's *beat~* Max/MSP object based on (Scheirer 1998), and density detection, where we look at the number of note onsets per time unit to represent the density of the rhythmic structure. We also developed a new implementation for higher-level rhythmic percepts of stability, based on (Desain, et al. 2002), and similarity based on (Tanguiane 1993). Desain's stability model is based on the relationship between pairs of adjacent note durations that are rated according to their perceptual expectancy. This depends on three main criteria: perfect integer relationships are favored, ratios have inherent expectancies (i.e., 1:2 is favored to 1:3 and 3:1 is favored to 1:3), and durations of 0.6 seconds are preferred. The expectancy function may be computed as:

$$E_b(A,B) = \int_0^r \left(\mathrm{round}(r) - r\right) \times \left|2\left(r - \mathrm{floor}(r) - 0.5\right)\right|^p \times \mathrm{round}(r)^d \, dr$$

where A and B are the durations of the two neighboring intervals, $r = max\ (A|B,B|A)$ represents the near-integer relationship between note durations, p controls the shape of the peaks, and d is negative and affects the decay rate as the ratios increases. This function is symmetric around $r = 1$ when the total duration is fixed. Generally, the expectancy function favors small near-integer ratios and becomes asymmetric when the total duration varies, exhibiting the bias toward the 600ms. interval. Our similarity rating is derived from Tanguiane's binary representation, where two rhythms are first quantized, and then given a score based on the number of note onset overlaps and near-overlaps. In order to support real-time interaction with human players, we developed two Max/MSP externals that analyze and generate rhythms based on these stability and similarity models. These externals have been embedded in a live interaction module that reads measure-length rhythmic phrases and modifies them based on desired stability and similarity parameters. Both parameters vary between 0 and 1 and are used together to select an appropriate rhythm from a database of pre-analyzed rhythms. A stability rating of 1 indicates the most stable rhythm in the database, 0.5 equates to the stability of the input rhythm, and 0 to the least stable rhythm. The similarity parameter determines the relative contribution of similarity and stability (see details in Section 5).

4.4 Interaction Design

Haile's original prototype incorporated three basic interaction modes – Imitation, Stochastic Transformation, and Simple Accompaniment. In the new design, we improved these modes, and added three advanced interaction schemes. In Imitation mode the original one-armed prototype repeated recorded rhythmic sequences based on its low-level onset, pitch, and amplitude perception modules. With its new two-arm design, the robot currently detects input rhythms in a similar manner but uses its left arm to play lower pitches and the right arm to play higher pitches, providing a more intuitive visual representation to players and audiences. In Stochastic Transformation mode, Haile currently improvises in a call-and-response manner based on players' input. Here, the robot stochastically divides, multiplies, or skips certain beats in the input rhythm, creating variations of users' rhythmic motifs while keeping their original feel. Different transformation coefficients can be adjusted manually or automated to control the level of similarity between input motifs and Haile's responses. In Simple Accompaniment mode the original one-arm prototype played pre-recorded MIDI files, allowing players to interact with it by playing their own rhythms or by modifying drumhead pressure to modulate and transform Haile's sound in real-time. In the current implementation, the Simple Accompaniment mode is designed for composers to feature their structured compositions in a manner that is not susceptible to algorithmic transformation or user input. This mode is usually used in sections of synchronized unisons where human players and Haile play together.

A number of new interaction modes – Perceptual Transformation, Beat Detection and Perceptual Accompaniment – were developed and added to Haile's interaction scheme. In Perceptual

Transformation mode, Haile analyzes the stability level of humans' rhythms, and responds by choosing and playing other rhythms that have similar levels of stability to the original input. In this mode Haile automatically responds after a specified phrase length. In Beat Detection mode, Haile utilizes the Max/MSP object *beat~* to track the tempo and beat of the input rhythm. The object beat~ was originally designed to analyze pre-recorded audio. In a real-time setting, however, human players tend to adjust to the robot's tempo, which leads to an unsteady input beat that is difficult for *beat~* to follow. Haile, therefore, uses *beat~* to listen for a short period (5-10 seconds) and then locks the tempo before joining in.

Perhaps the most advanced mode of interaction is the Perceptual Accompaniment mode, which combines synchronous, sequential, centralized and decentralized operations. Here, Haile plays simultaneously with humans while listening to and analyzing their input. It then creates local call-and-response interactions with different players, based on its perceptual analysis. In this mode we utilize the amplitude and density perceptual modules described above. While Haile plays short looped sequences (captured during the Imitation and Stochastic Transformation modes) it also analyzes the amplitude and density curves of human playing. It then modifies its looped sequence, based on the amplitude and density coefficients of the human players. When the rhythmic input from the human players is dense, Haile plays sparsely, providing only the strong beats and allowing humans to perform denser solos. When humans play sparsely, on the other hand, Haile improvises using dense rhythms that are based on stochastic and perceptual transformations. Haile also responds in direct relationship to the amplitude of the human players so that the louder humans play the stronger Haile plays and vice versa, accommodating its drumming to humans' dynamics[1].

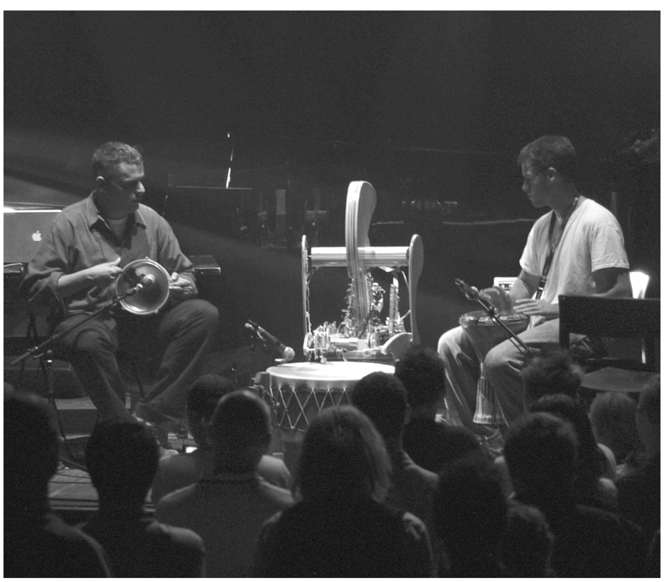

Figure 6. Human-robot interaction with Haile in a performance ("RoboRave" Odense, Denmark, September 2006)

[1] See a video clip of several interaction modes in performance: http://coa.gatech.edu/~gil/Jam'aaJerusalem.mov

5. USER STUDY

5.1 Method

In order to evaluate our approaches in design, mechanics, perception and interaction we conducted a user study where subjects were asked to interact with Haile, to participate in a perceptual experiment, and to fill a questionnaire regarding their experience. The 14 undergraduate students who participated in the study were enrolled in the percussive ensemble class at Georgia Tech in Spring 2006 and had at least 8 years of experience each in playing percussion instruments. This level of experience was required to facilitate the musical interaction with Haile as well as to support a meaningful discussion about subjects' experience. Each subject spent about 20 minutes experimenting with four different interaction modes – imitation, stochastic transformation, perceptual accompaniment, and perceptual transformation. As part of the perceptual experiment on stability, subjects were asked to improvise a one-measure rhythmic phrase while Haile provided a 4/4 metronome beat at 90 beats per minute. Subjects were then randomly presented with three transformations of their phrase: a less stable, similar stability, and more stable version. The transformed measures were generated by our Max/MSP stability external (see section 4.3) using stability ratings of 0.1, 0.5, and 0.9 for less, similar, and more stability, respectively. All phrases, including the original, were played twice. Students were then asked to indicate which phrase, in their opinion, was less stable, similar, or more stable in comparison to the original input. Stability was explained as representing the "predictability of" or "ease of tapping one's foot along with" a particular rhythm. The goal of this experiment was not to reach a definite well-controlled conclusion regarding the rhythmic stability model we used, but rather to obtain a preliminary notion about the correlation between our algorithmic implementation and a number of human subjects' perception in an interactive setting[2].

The next section of the user study involved a written survey where subjects were asked to answer questions describing their impression of Haile's physical design, mechanical operation, the different perceptual and interaction modes, as well as a number of general questions about human-robot interaction. The survey included 39 questions such as: "What aspects of the design and mechanical operation make Haile compelling to play with?" "What design aspects are problematic and require improvements"? "What musical aspects were captured by Haile in a satisfactory manner"? "What aspects were not captured well?" "Did Haile's response make musical sense?" "Did the responses encourage you to play differently than usual and in what ways?" "Did the interaction with Haile encourage you to come up with new musical ideas?" "Do you think that new musical experiences, and new music, can evolve from musical human-robot interaction? "

5.2 Results

Most subjects addressed Haile's physical design in positive terms, using descriptors such as "unique", "artistic", "stylized", "organic", and "functional". Other opinions included "the design offered a feeling of comfort", "the design was pleasing and inviting, and "if Haile was not anthropomorphic it would not have been as encouraging to play with". When asked about caveats in the design

[2] See a short video clip from the user study: http://coa.gatech.edu/~gil/HaileUserStudy.mov

several subjects mentioned “too many visible electronics” and “exposed cabling” and suggested that future designs should be “less cluttered.” Another critique was that “the design did not appear to be versatile for use with other varieties of drums.” Regarding Haile’s mechanical operation, subjects provided positive comments about the steadiness and accuracy of the left hand and the speed and “smoothness” of the right hand. The main mechanical caveats mentioned were Haile’s limited timbre and volume control as well as the lack of larger and more visual movements in the right hand. Only one respondent complained about the mechanical noise produced by Haile. In the perceptual rhythmic stability study, half of the respondents (7 of 14) correctly identified the three transformations (in comparison, a random response would result in 2.3 of 14 correct choices on average). The majority of confusions were between similar and more stable transformations and between similar and less stable transformations. Only 3 responses out of the total 42 decisions confused a more stable version for a less stable version, implying that larger differences in algorithmic stability ratings made differentiation easier. Only one subject labeled all three generated rhythms incorrectly.

Subjects’ response to the four interaction modes was varied. In Imitation Mode respondents mentioned Haile’s “accuracy,” and “good timing and speed” as positive traits and its lack of volume control as a caveat. Responses to the question “How well did Haile imitate your playing?” ranged from “pretty well” to “amazingly well.” Some differences between the interaction modes became apparent. For example, in Stochastic Transformation Mode (STM), about 85% of the subjects provided a clear positive response to the question “Was Haile responsive to your playing?” Only about 40% gave such a clear positive response to this question in Perceptual Accompaniment Mode (PAM). Respondents refer to the delay between user input and robotic response in PAM as the main cause for the “less responsive feel.” To the question “Did Haile’s responses encourage you to play differently than usual?” 50% of the subjects provided a positive response in STM while only 30% gave a positive response to this question in PAM. When asked to describe how different than usual their playing in STM was, subjects focused on two contradicting motivations: Some mentioned that they played simpler rhythms than usual so Haile could transform them easily and in an identifiable manner. Others made an effort to play complex rhythms to challenge and test Haile’s abilities. These behaviors were less apparent in PAM. While only 40% (across all interaction modes) provided a positive answer to the question “Did Haile’s responses encourage you to come up with new musical ideas”?, more than 90% percent of participants answered positively to the question “Do you think that new musical experiences and new music, can evolve from human-machine musical interactions?”, strengthening their answers with terms such as “definitely,” “certainly,” and “without a question”.

6. DISCUSSION

Based on the experiment and survey, we feel that the second phase of the project demonstrated significant improvements in the musical effectiveness of the robot-human interaction in comparison to our original prototype. The most encouraging survey outcome, in our opinion, was that subjects felt that the collaboration established with Haile did, on occasions, lead to novel musical experiences and new musical ideas that would not have been conceived by other means. It is clear, though, that further work in mechanics, perception, and interaction design is required to create a robot that can demonstrate comprehensive musicianship. Nearly all subjects addressed Haile’s design in positive terms, strengthening our assumption that the wood and the organic look function well in a musical context. Our decision to complement the organic look with exposed electronics was criticized by some subjects, although we feel that this hybrid design conveys the robotic functions and reflects the electroacoustic nature of the project. Mechanically, most subjects were impressed with the speed and smoothness of Haile’s operation. Only one subject complained about the noise produced by the robot, which suggests that most players were able to either mask the noise or to accept it as an inherent and acceptable aspect of human-robot interaction. Several subjects indicated that Haile’s right arm’s motion did not provide satisfactory visual cues and could not produce adequate variety of loudness and timbre. We, therefore, consider changing Haile’s right arm design to utilize a linear motor similarly to its left arm. In addition, we intend to further improve the control for the left arm in an effort to increase timbre variety though various techniques such as pressing on the skin while the other hand plays and applying damping briefly after hits (see Future Work section).

The user study and survey provided encouraging results in regard to Haile’s perceptual modules. The high percentage of positive responses about the Imitation Mode indicates that our low-level onset and pitch detection algorithms were accurate and effective. In general, a large majority of the respondents indicated that Haile was responsive to their playing. Perceptual Accompaniment Mode (PAM), however, was an exception to this rule, as subjects felt Haile was not responding to their actions with acceptable timing. PAM was also unique in the high percentage of subjects who reported that they did not play differently in comparison to playing with humans. We explain this results by the synchronous accompaniment nature of PAM, which is more familiar to most percussion students. Most subjects, on the other hand, felt compelled to play differently than usual in sequential call-and-response modes such as Stochastic Transformation Mode (STM). Here subjects changed their usual drumming behaviors either by simplifying their rhythms to better follow Haile’s responses or by playing complex rhythms in an effort to challenge the robot’s perceptual and mechanical abilities. We believe that these behaviors were the results of players’ attempt to explore and accommodate to Haile’s physical and cognitive boundaries. We assume that subjects may develop more complex interaction behaviors if given longer play times and more advanced interaction schemes.

The results of our rhythmic stability experiment were mixed. However, given the high level of variance in the notion of rhythmic stability in human perception we feel that the algorithm performed better than expected. Some caveats in our method may have also hindered the results. For example, misalignment of subject drumming with the metronome during recording led to misaligned transformations, which may have been perceived as unstable. Also, since the transformed rhythms were generated based on subjects’ input, the relative difference between the output stabilities in some cases became minimal and difficult to identify. For example, when a subject’s original phrase was extremely stable the algorithm would not be able to produce an identifiably “more stable” phrase. Asking subjects to play a unified mid-stability rhythm as input could have solved this problem, although we were specifically interested in evaluating Haile’s perception in a live improvisatory context.

7. FUTURE WORK

We plan to pursue several research directions in immediate and long-term future work. Mechanically, we intend to further investigate hand drumming in an effort to improve Haile's timbre control. The sound quality of a drum is dependent on a large number of subtle factors (Taylor 2004). A wide sonic variation can be produced by playing a hand drum using different stroke motions, contact areas (fingers vs. palm, etc.), leaving the hand momentarily on the skin to briefly dampen it, stretching the skin with the other hand, etc. Most mechanical instruments (player pianos, drums, mallets, etc.) only produce quick bounces off the surface and avoid "human finesse" during the hit. But it is this finesse that makes a human player's expression and intonation interesting and colorful. We believe that current state-of-the-art technology cannot support the creation of a robot with such dexterity that would compare with humans' expressivity and virtuosity, but we do believe that significant advancements are possible. Furthermore, we would also like to expand on what is possible by humans by experimenting with alternative striker shapes, materials, and mechanisms that do not necessarily reflect human organs and techniques. In light of our plan to continue improving Haile's mechanical stroke variety and timbre control, we also plan to explore better approaches for the perception of timbre and stroke variety. To this end we are examining a number of neural network and machine learning approaches such as in (Tindale, et al. 2005) and (Chordia 2004).

In the longer term, we also hope to expand the promise of robotic musicianship to pitch mallet-based instruments (such as the vibraphone or the marimba). This direction may also call for further research into perceptual modeling and interaction design with pitch and tonality, allowing Haile to listen and respond to pitch based monophonic and polyphonic input. Some of the percepts that we have started to investigate in that regard are local attractions, melodic similarity (Hewlett, et al. 1998), and melodic complexity (Narmour 1992). We also consider implementing models for melodic attraction (Lerdahl, et al. 1983), melodic tension (Narmour 1990), and contour directionality (Trehub, et al. 1984). The choice of percepts and modeling schemes will be integral to the definition of future interaction and response algorithms. We plan to continue evaluating our current interaction design with more in depth user studies and to improve and fit future interaction modes to Haile's new capabilities. In particular, we are interested in designing new interaction schemes that would take advantage of Haile's ability to listen to and interact with multiple players in a group.

In addition to expanding our research in mechanics, perception, and interaction design, we also plan to investigate the use of Haile for educational purposes. Our educational pedagogy is informed by the theory of Constructionism, which demonstrates how learning is most effective when students construct personally meaningful technological artifacts (Papert 1980). The theory emphasizes the unique ability of digital technology to provide personal and configurable learning experiences to a wide variety of learners. More recent research elaborated on Papert's ideas, showing how interaction with digital physical objects enhances children's and adults' learning (Resnick, et al. 1996). In the field of music, however, little has been done to develop physical constructionist systems that can provide a compelling interdisciplinary education, not only in music, but also in math, sciences, and programming. For our educational work with Haile, therefore, we hope to capitalize on the beneficial effect of music education on learning in domains of knowledge beyond that of music (Rauscher 1993) (Schillinger 1976) (Bamberger 2000). We plan to build on our previous work in this area (Weinberg, et al. 2000) by developing a constructionist educational application for Haile that would allow learners to translate abstract musical ideas into symbolic representations and physical gestures. The mathematical and scientific aspects of the project will be guided and motivated by learners' drive to rhythmically compose acoustic sounds, creating personal interactive musical compositions. The educational environment will allow learners to develop their intuitions by emphasizing shared structures in music and math such as hierarchies, periodicity, units, ratios and proportions, symmetries, and patterns. Students will also be able to experiment with creating perceptual social behaviors by programming rule-based responses in an effort to make Haile an expressive, responsive, and intriguing playing companion.

8. ACKNOWLEDGMENTS

We would like to thank Georgia Tech's GVU Center, the College of Architecture and the Music Department, who supported this project.

9. REFERENCES

[1] Baginsky, N.A.. The Three Sirens: A Self-Learning Robotic Rock Band. *http://www.the-three-sirens.info*. Accessed July 2006.

[2] Bamberger, J. Music, Math, and Science: Towards an Integrated Curriculum. *Journal for Learning Through Music, (*Spring 2000), 32-35.

[3] Chordia, P. 2004. Automatic Tabla Transcription. PhD Thesis Stanford University.

[4] Coren, S. and Ward, L.M.. *Sensation and Perception*. Harcourt Brace Jovanovich, San Diego, 1984.

[5] Coyle, E.J. and Shmulevich, I.A. System for Machine Recognition of Music Patterns. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*. Seattle, WA ,1998, 3597-3600.

[6] Cycling74. Max/MSP. *http://www.cycling74.com* Accessed July 2006.

[7] Dannenberg, R. An On-line Algorithm for Real-Time Accompaniment. In *Proceedings of the International Computer Music Conference* Paris, France, 193-198.

[8] Dannenberg, R.B., Brown, B., Zeglin, G., and Lupish, R. McBlare - A Robotic Bagpipe Player. In *Proceedings of the International Conference on New Interface for Musical Expression*. Vancouver, Canada, 2005. 80-84.

[9] Desain, P. and Honing, H.J. Rhythmic stability as explanation of category size. In *Proceedings of the International Conference on Music Perception and Cognition* ,Sydney, 2002 CD-Rom.

[10] Foote, J. and Uchihashi, S. The Beat Spectrum: a new approach to rhythmic analysis. In *Proceedings of the International Conference on Multimedia and Expo*, Tokyo, Japan, 2001, 881- 884.

[11] Hewlett, B. and Selfreidge-Field, E. *Melodic Similarity - Concepts, Procedures, and Applications*. MIT Press. Cambridge, MA 1998.

[12] Jordà, S. Afasia: The Ultimate Homeric One-man multimedia-band.. In *Proceedings of the The International Conference on New Interfaces for Musical Expression*. Dublin, Ireland, May 24-26 2002, 1-6.

[13] Kapur, A. A History of Robotic Musical Instruments. In *Proceedings of the 2005 International Computer Music Conference*, Barcelona, Spain ,21-28.

[14] Lerdahl, F. and Jackendoff, R. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA, 1983.

[15] Lewis, G. Too Many Notes: Computers, Complexity and Culture in Voyager *Leonardo Music Journal* 10 Cambridge, MA, (2000) 33-39

[16] MakingThings. Teleo. *http://www.makingthings.com* . Accessed July 2006

[17] Narmour, E. *The Analysis and Cognition of Basic Melodic Structures.* Chicago University of Chicago Press, 1990

[18] Narmour, E. T*he Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. Chicago University of Chicago Pres, 1992

[19] Orio, N., Lemouton, S., Schwarz, D. Score Following: State of the Art and New Developments. In *Proceedings of the International Conference on New Interfaces for Musical Expression.* Montreal, Canada, 1992, 36-41.

[20] Pachet, F. The continuator: Musical interaction with style. In *Proceedings of the International Computer Music Conference*. Göteborg, Sweden, 2002, 211-218

[21] Papert, S. *Mindstorms*. New York Basic Books, 1980

[22] Paulus, J. and Kalpuri, A. Measuring the Similarity of Rhythmic Patterns. In *Proceedings of the 2002 International Conference on Music Information Retrieval*. Paris, France, 150-156

[23] Puckette, M. Real-time Audio Analysis Tools for Pd and MSP. In *Proceedings of the International Computer Music Conference*, Ann Arbor, MI, 1998, 109-112

[24] Rae, G.W. Robotic Instruments. *http://logosfoundation.org/instrum_gwr/automatons.html* Accessed July 2006

[25] Rauscher, F.H, Shaw, G.L, and Ky, C.N. 1993. Music and spatial task performance *Nature* 365, 611

[26] Resnick, M., Martin, F., Sargent, R., and Silverman, B. 1996. Programmable Bricks: Toys to Think With. *IBM Systems Journal*, 35:3&4

[27] Rowe, R. 2004. *Machine Musicianship*. Cambridge, MA MIT Press.

[28] Scheirer, E. 1998. Tempo and beat analysis of acoustic musical signals *Journal of the Acoustical Society of America*, 103:1 588-601

[29] Schillinger, J. *The Mathematical Basis of the Arts*. Da Capo Press, New York, 1976.

[30] Singer, E., Feddersen, J., C., R., and Bowen, B. LEMUR's Musical Robots. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Hamamatsu, Japan, 2004

[31] Singer, E., Larke, K., and Bianciardi, D. LEMUR GuitarBot: MIDI Robotic String Instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Montreal, Canada, 2003

[32] Sony. 2003. QRIO conductor Robot. *http://www.sony.net/SonyInfo/QRIO/works/20040325e.html* Accessed July 2007

[33] Takanishi, A. and Maeda, M. Development of Anthropomorphic Flutist Robot WF-3RIV. *In the Proceedings of International Computer Music Conference*, 1998 , 328-331.

[34] Tanguiane, A. *Artificial Perception and Music Recognition.* Springer-Verlag, New York,1993

[35] Taylor, M. *Remembering How to Drum: Djembe Technique* (DVD). Blue 7 Media, 2004

[36] Tindale, A.R., Kapur, A., Schloss, W.A., and Tzanetakis, G. Indirect Acquisition of Percussion Gestures Using Timbre Recognition. In *Proceedings of the Colloque Interdisciplinaire de Musicologie*. Montreal, 2005

[37] Toyota. 2004. Trumpet Robot. *http://www.toyota.co.jp/en/special/robot* Accessed July 2006

[38] Trehub, S.E., Bull, D., and Thorpe, L.A. 1984. Infants' perception of melodies: The role of melodic contour *Child Development*, 55:1 821-830

[39] Vercoe, B. The Synthetic Performer in the Context of Live Performance. In *Proceedings of the International Computer Music Conference*. Paris, France, 1984 199-200.

[40] Weinberg, G. 2005. Interconnected Musical Networks - Toward a Theoretical Framework, *Computer Music Journal 29:2*, MIT Press, Cambridge, MA, 23-39

[41] Weinberg, G., Lackner, T., and Jay, J. The Musical Fireflies - Learning About Mathematical Patterns in Music Through Expression and Play. In P*roceedings of the Colloquium on Musical Informatics* 2001. A'quila, Italy 146-149.

[42] Weinberg, G., Driscoll S., and Parry, M. Musical Interactions with a Perceptual Robotic Percussionist. In *Proceedings of the of IEEE International Workshop on Robot and Human Interactive Communication,* 2005, Nashville, TN 456-461.

[43] Weinberg, G., Driscoll, S., and Thatcher, T. Jam'aa for Haile. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRPAPH 2006)*. 2006 Boston MA. p. 124.

[44] Winkler, T. *Composing Interactive Music: Techniques and Ideas Using Max.* Cambridge, MA MIT Press, 2001