Dancing cheek to cheek: Haptic communication between partner dancers and swing as a finite state machine

by

Sommer Elizabeth Gentry

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy of Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2005

Author . . . .
Department of Electrical Engineering and Computer Science
May 18, 2005

Certified by . . . .
Eric Feron
Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Dancing cheek to cheek: Haptic communication between partner dancers and swing as a finite state machine

by

Sommer Elizabeth Gentry

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy of Electrical Engineering and Computer Science

Abstract

To see two expert partners, one leading and the other following, swing dance together is to watch a remarkable two-agent communication and control system in action. Even blindfolded, the follower can decode the leader's moves from haptic cues. The leader composes the dance from the vocabulary of known moves so as to complement the music he is dancing to. Systematically addressing questions about partner dance communication is of scientific interest and could improve human-robotic interaction, and imitating the leader's choreographic skill is an engineering problem with applications beyond the dance domain.

Swing dance choreography is a finite state machine, with moves that transition between a small number of poses. Two automated choreographers are presented. One uses an optimization and randomization scheme to compose dances by a sequence of shortest path problems, with edge lengths measuring the dissimilarity of dance moves to each bar of music. The other solves a two-player zero-sum game between the choreographer and a judge. Choosing moves at random from among moves that are good enough is rational under the game model. Further, experiments presenting conflicting musical environments to two partners demonstrate that although musical expression clearly guides the leader's choice of moves, the follower need not hear the same music to properly decode the leader's signals.

Dancers embody gentle interaction, in which each participant extends the capabilities of the other, and their cooperation is facilitated by a shared understanding of the motions to be performed. To demonstrate that followers use their understanding of the move vocabulary to interact better with their leaders, an experiment paired a haptic robot leader with human followers in a haptically cued dance to a swing music soundtrack. The subjects' performance differed significantly between instances when the subjects could determine which move was being led and instances when the subjects could not determine what the next move would be. Also, two-person teams that cooperated haptically to perform cyclical aiming tasks showed improvements in the Fitts' law or Schmidt's law speed-accuracy tradeoff consistent with a novel endpoint compromise hypothesis about haptic collaboration.

Thesis Supervisor: Eric Feron
Title: Associate Professor of Aeronautics and Astronautics

Acknowledgments

First, I must acknowledge the support of my wonderful fellowship program, the U.S. Department of Energy's Computational Science Graduate Fellowship (CSGF). Through summertime conferences, generous travel grants, mentorship from computational science leaders, a Sandia Laboratories practicum, and best of all the cameraderie of the large group of fellows; the CSGF program shaped my success.

I am very grateful to the Lindy Hop community, in which I have found much more than a research topic. I've found a true home. The genuine friendship that the dance community extends to all of its members, even those we haven't met yet, is second to none. Thank you to all the dancers who wore spandex and velcro dots, or who wore headphones, or who wore blindfolds, in the interest of the research documented here. Your generosity made it possible.

Thank you to Dava Newman, an exemplar of fine teaching, in whose Space Biomedical Engineering class I first attempted to analyze swing dancing biomechanics quantitatively.

Thank you to Roderick Murray-Smith for his careful advising at a critical point in my work. Thank you also to the Hamilton Institute of the National University of Ireland, Maynooth, for hosting me during my stay in Ireland.

Thank you to Allison Okamura for welcoming me to her lab at Johns Hopkins University, and for her detailed and rigorous editing of the final versions of my haptics chapters.

This thesis would never have come about if not for my advisor, Eric Feron. He encouraged me to study what I was passionate about, whether or not that fit within neat disciplinary boundaries. He offered me the freedom to take my own path through graduate school, and I hope that he will be pleased with the path I chose.

Thank you to Sarah Smyth and John Smyth, who kept me off the streets.

To my parents, Larry and Carole Gentry, I offer in return the gift you gave me. To know that one is deeply loved is the most valuable gift any person can give another. Mom, let's go to Vegas. Dad, let's go climb a hill like mountain goats.

Dorry, you are the partner I will always need. It was a stroke of incredible luck for me to find in one person a dance partner, a research partner, a teaching partner, a musical partner, and a life partner. Let's just keep building amazing things.

I attribute my small achievements to the people mentioned above, but I keep the burden of any weaknesses in this manuscript for my own.

Biographical Note

Sommer Gentry earned a B.S. in Mathematical and Computational Science and an M.S. in Operations Research from Stanford University in 1998. Before beginning her doctorate, she was an Engineer in the Systems Sciences division at Lawrence Livermore National Laboratories. She received the Best Student Paper award at the 2003 IEEE Conference on Systems, Man, and Cybernetics. She has been a guest on the Diane Rehm show on National Public Radio, and her work has been mentioned in Science, The Boston Globe, and 5 6 7 8 Dance magazine. Starting August 2005, she will be Assistant Professor in the Mathematics Department at the United States Naval Academy.

Sommer is also an avid Lindy Hop dancer with seven years of experience teaching and competing in this historical swing dance form. She teaches Lindy Hop at workshops worldwide, and has taught dancers in Albuquerque, Amsterdam, Baltimore, Boston, Dublin, Edinburgh, London, Portland (Maine), and Washington, DC. She and her husband, Dr. Dorry Segev, have placed in the American Lindy Hop Championships three times. They won the 2002 United Kingdom Lindy Hop Championships and have won or placed in dozens of other major swing dance competitions. They have performed as featured dancers with the Baltimore Symphony Orchestra and at the Allentown Symphony Hall and Boston's Wang Center. They also teach dance online and via a series of swing dance CD-ROMs.

Contents

List of Figures

List of Tables

Chapter 1

Swing dancing: what's there to know?

> All the ills of mankind, all the tragic misfortunes that fill the history books, all the political blunders, all the failures of the great leaders have arisen merely from a lack of skill at dancing.
>
> -Molière

1.1 And a 1, and a 2: all about swing dance

Millions of people have been dancing to American swing music, for the last seventy-five years, all over the world. They love the infectious beat and the playful freedom of the dances called Lindy Hop, Jitterbug, and just plain Swing. But to enthusiasts, partner dancing is much more than good exercise. To the swing dance fanatics who have invested years of study and practice, each dance is a conversation. The leader takes pride in choosing the moves that express the subtleties of a song, while the follower is cheered at each complicated movement that she can execute properly because she correctly read the leader's intent.[1]

While a solo dancer may choose freely from the universe of movement, limited only by the dancer's own ability to enact it, the necessity of communication in partner dancing does constrain its adherents in important ways. In lead and follow partner dances, one partner

[1] Leaders are not always male, nor are followers always female. In fact, there are talented male followers and female leaders in competitive swing dancing. This text will label leaders as male and followers as female for the sake of brevity and simplicity.

indicates the sequence of moves while the other partner deciphers the signals that tell what move will come next. Square dancing or ballet partnering are not lead and follow dances, since in square dancing the moves are called out to all participants and in ballet partnering the moves are known in advance.

In swing dancing, the only indispensable elements for lead and follow are rhythm and touch. Touch makes a difficult channel over which to send any sort of message for many reasons: force perception and proprioception are notoriously inaccurate, forces imposed on the follower can easily unbalance her, and kinematic redundancy means there are many different ways that imposing a certain force might move a person's body. This introduction will examine what all partner dances share: the elements that enable communication between the leaders and followers. Throughout the discussion, the primary dance considered is Lindy Hop or swing dance [8, 13]. Following this dancer's description of swing, Section 1.2 will review some relevant literature and Section 1.3 describe the contributions of the dissertation.

1.1.1 Rhythm

The rhythm of a dance is the timing of the motion, its synchronization with the music. In simple terms, each foot should hit the ground at the same time that the drums or bass or piano hit the beat. The best dancers usually hit closer to the end of each beat than to the beginning or the middle, because that makes their dancing look relaxed and natural, where hitting in the early portion of the beat can make dancing look rushed or mechanical.

Both the leader and the follower can hear the beat and each can entrain his movements to the external metronome. When people claim to have no rhythm, they probably mean that they have difficulty entraining to an external metronome, or that they have trouble hearing where the metronome element is when listening to a song. The common rhythm between two dancers is the strongest element tying their actions together. The rhythm not only lets both leaders and followers predict when the current pattern will end, but it allows each partner to accurately gauge when the other partner plans to step or move. In fact, dancers who hear music with beats mismatched by the smallest fraction of the interbeat time will soon find themselves unable to continue.

In lindy hop, the original form of swing dance, the basic rhythm goes 1, 2, 3-and-4, 5, 6, 7-and-8. The 3-and-4 section is syncopated, so that the "and" will be slightly closer in time to the 4 than it was to the 3. Try tapping this on a desk to see if you can syncopate your tapping. Every count in the rhythm corresponds to one step with a weight change in the basic footwork, so the basic footwork has ten steps per eight beats.

Swing music and swing dance emphasize the downbeats, which are the 2, 4, 6, 8 of the above sequence. In fact, a primary chronicler of swing music in its heyday was the jazz magazine *Down Beat*, which is still published today. A subtle form of music misinterpretation is to know when the beats are, but to be unable to distinguish the upbeats from the downbeats. This will very occasionally happen to swing dancers, and the result is one partner dancing the 1 2 3 4 of a move while the other dances the 2 3 4 5 of that move. This problem is so disruptive that the pair will have to stop dancing to agree on the downbeat before they start moving again. If the dancers are off by two beats from each other, then they are back in agreement about which is the downbeat, and it doesn't matter very much whether a section of music is the 1 2 3 4 or the 3 4 5 6. It is true that many dancers try to begin an eight-count pattern on the 1. In most music, the 1 sounds more emphasized than the 3 or the 7. However, this alignment is not as essential as the upbeat - downbeat alignment.

In fact, a bar of swing music is generally four beats in length, so for a musician there are no numbers beyond 4. "And a 5 6 7 8" is a dancer's phrase, created to count through longer moves that cover two bars of music. We will use the term *bar* to mean a dancer's bar, which is 8 counts long.

1.1.2 Positions or poses

An extremely simple dance might have only a single pose. For instance, a couple could dance an entire waltz without ever changing from the pose or position that they came together in, the closed position. Figure 1-1 shows a couple in closed position. Closed position is: with about a 100° angle between the partners, the leader's right hand is around the follower's back and the leader's left hand holding the follower's right.

Any lindy hop dance will have multiple poses or positions, because the basic move of lindy hop includes a transition from open position to closed position and back to open position. Frequent switching between positions is one of the elements that makes lindy hop a challenging dance. Lindy hop moves use a very few positions. Excepting a few variants, all of the positions the partners reach will be either closed or open. Figure 1-2 shows a

Figure 1-1: Closed position

couple in open position. Open position is: with partners facing each other from about three feet away, the leader's left hand is holding the follower's right. Each dancer should have his elbow bent to about a 150° angle and should keep the arm bent throughout the dance to avoid the singularity that results from a completely straight arm. Sometimes open position has the leader's right hand holding the follower's left, which is called cross hand, or the leader's right hand holding the follower's left, which is called wrong hand.

Dance moves begin and end from one of these fixed positions. This rule allows the follower to arrive at the place the leader expects her (in either closed or open position) at the time when he expects her to be there (at the end of the bar, on the downbeat). For instance, if a dancer leads his partner to spin clockwise and then lets go of her, she will know that she should end her spin facing him because open position faces him. As she spins, he will position himself roughly the same three feet away that is standard in open position. Even if her turn wobbles across the floor, he can track her movement so that the couple ends the spin in the ideal open position.

1.1.3 Choreography

Each partner dance has a small core of patterns, called moves, that all dancers of that form are exposed to. Just as square dancing has its do-si-do, Lindy hop has the move called a swingout. A swingout, also referred to as a whip, is an 8-beat move in which the leader and the follower start in open position. They approach each other, reach closed position,

Figure 1-2: Open position

rotate a full 360° around each other, and then return to an open position. There are endless variations on this theme. For example, either the leader or the follower can do extra spins, in either direction, before or after the closed position rotation.

Having a small set of known moves helps the two dancers coordinate their actions. Between a leader who knew the moves and a follower who did not, the dance would not be very successful, although an expert leader could communicate some of the requested motion to a naive follower. Chapter 5 demonstrates how people can respond more nimbly when given a specific vocabulary of known moves, in a series of experiments using a simple robotic leader with a human follower. Lindy hop also has more general “traffic rules”, like the rule that followers keep moving in the same direction until something makes them stop. Such rules enable leaders to send experienced followers through moves the follower has never seen before.

Followers have multiple redundant mechanisms for decoding the leaders’ signals and deciding which move he is requesting. Since dance is intended to elaborate upon music through movement, the shared musical environment can indicate whether the next move should be exciting or mellow, sharp or fluid. This influence of the emotional content of music is distinct from the rhythmic, metronome input described in section 1.1.1. Sometimes the music is called “the third partner” in that it may suggest a leader’s likely choice of moves. Surprisingly, the communication between leaders and followers is robust to disturbances in the musical meaning of songs, as demonstrated by experiment in Chapter 2.

Assigning meaning to the songs and dance moves is a quintessentially human task. The impression of similarity between musical expression and the dance moves that accompany it is called musicality. After human judges translate their impressions to numerical scores, a computer can run an optimization algorithm to choose a move sequence with musicality. Two automated choreographers of this type are described in the thesis: one based on a shortest-path algorithm with randomization appears in Chapter 3, and a second based on a two-person zero-sum game model of the interaction between dancers and judges appears in Chapter 4.

The leader chooses a sequence of moves to satisfy several goals. While a very small and predictable set of moves may make coordination with his follower easier, a surprising and novel sequence from a larger set of moves will be more interesting to an audience or contest judge. In Chapter 4, studies of the dance sequences chosen in national lindy hop contests reveal that the leader favors surprise when dancing a pre-arranged routine with this regular partner, and favors predictability when dancing improvised dances with a random partner.

1.1.4 Connection

In open position, the contact between the leader and follower is maintained only at a single point: the leader's left hand holding the follower's right hand. Messages and energy are transmitted through the connected hands. For instance, if the leader holds the follower's hand and takes a step backward, then the leader's hand has moved back and thus the follower's hand will move forward. An expert follower will not let her hand move very much relative to her body position, and so because her hand has moved forward she will take a step forward. An inexperienced follower might let her arm extend while she stays in place, in which case she has lost the opportunity to use the leader's energy to translate her center of mass. The inexperienced follower will now have to use her own energy to move forward and return to the typical open position. In either case, some message is transmitted through the connection hand. This is in essence how leading happens. Expert followers and leaders perform very well when deprived of visual cues [blindfolddance.mpg][2], so the physical connection of hands is a sufficient means of communication between the partners.

[2] An accompanying CD contains a number of video and music files. Where each file is referred to in the text, the filename will appear in brackets like this [filename.mpg]. A complete table of the multimedia files on the CD appears as an appendix.

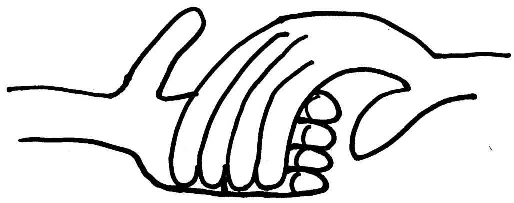

Figure 1-3: Connection: contact without gripping

Ordinarily, gripping an object guarantees that contact can be maintained. That is, to hold an apple or a piece of paper, a person applies force to one side of the object with the thumb and to the other side of the object with some fingers. In dancing, one partner does not grip the other's hand, and not just because being gripped tightly is uncomfortable. Mostly, gripping is forbidden because both dancers turn frequently and so one hand needs to rotate freely within the other. There is an expression, "all thumbs", describing a clumsy person, so it is perhaps apt that graceful dancers use no thumbs at all.

Instead of gripping, dancers use a strategy of imposing slight forces on each other's hands in the direction that pushes the hands together, as in Figure 1-3. This is the same as a person's two handed strategy for lifting a box without gripping, described by Zefran and colleagues [144]. Even cooperating robots can use this strategy for non-prehensile lifting [30]. As you can see in the diagram, each partner needs constantly to create a force towards his or her body in order for the four fingers of the hands to be pressed together. The partners may do this by leaning very slightly away from each other. The force required is fairly small, so the difference from each partner balancing on his own is almost imperceptible visually. However, if dancers in this position were to suddenly separate their hands, each should have to take a step backwards. In this way, the dancers achieve a shared balance that is, surprisingly, not at all precarious. If the dancers do not create a force to push their hands together, then no force will keep his hand touching hers, and so when he moves his hand she may lose track of it.

There are some swing moves, called breakaways, during which the leader lets go of the follower. Breakaways are generally very short and the dancers regain contact and shared balance as each move ends. The rhythm which both dancers hear indicates when the contact will be regained, while their knowledge of the basic positions assures that they will come together in a position that transitions to the next move.

Figure 1-4: Upright stance

Figure 1-5: Laid out stance

1.1.5 Technique

Technique means both the way in which the moves can best be executed and the general traffic rules alluded to earlier. Dancers might not need advanced technique to lead a small set of moves to a slower song. However, across many different dances and styles, a large subset of technique remains valid. The expert lindy hop dancers in Figures 1-4 and 1-5 are both pictured at the same instant of a swingout. While they make hugely different visual impressions, they actually maintain very similar forces between their connection hands. One school of thought among dancers holds that the physics of doing a swingout at various speeds requires connection forces within a narrow acceptable range.

Probably the most fundamental point of technique is that a leader initiates motion and the follower responds to motion using translations and rotations of the person's center of mass. A follower who lets her arms travel widely with no relation to the travel of her center of mass is probably following poorly, and this technique problem is called spaghetti arms. A leader who requests motion from his follower without moving his own center of mass is probably leading poorly, and this is called an arm lead to contrast it with what would be better, a body lead.

Given that where to put the body's center of mass is the primary concern in partner dancing, it is unfortunate that many dance teachers describe moves by telling students where to put their feet. Where a dancer puts his feet should be determined entirely by where he puts his body, that is, if he commits his weight to the foot he is stepping on. Putting your left foot on the ground twelve inches forward without moving your center of gravity twelve inches forward creates no end of trouble in dancing.

All of these are the observations of a community of experienced partner dancers about balance, technique, and travel. However, introspection is a famously bad method for discovering facts about human motion. How much of this discussion could be made rigorous in the scientific sense? First of all, a scientist would need to determine whether the description given by dancers about proper technique in fact matches what expert dancers do. Second, a scientist could try to explain dance technique by reference to any number of engineering concepts: stability analysis to ensure the dancers stay upright, information theory to encode and decode move signals given the unreliable touch channel, optimization to minimize energy expenditure while executing a specific move together. Finally, to demonstrate that a complete understanding of partner dance has been achieved, a Turing test for the skill of partner dancing would require that an artificially constructed partner convincingly dances with a human partner.

It may be years or decades before a constructed robot can imitate the skills and competencies of a trained dancer. Even computed simulation of a complex interaction like lead and follow dancing would require intensive hand-tuning to achieve the fidelity that would make it useful. We have not taken that route, of quantitative dynamical analysis, in this work. Rather, we have concentrated on proving simple statements about how haptic communication between partner dancers functions, and on imitating choreography, the high-level planning of move sequences to harmonize the expressive content of the dance with the expressive content of the music.

1.2 Related work, applications

We have discussed at length in Section 1.1 the observations of a community of experienced partner dancers about the skills that enable their dancing. Even with the weight of introspection and experience, at least some of the canards about dancing that dancers embrace might contradict available evidence about human motion, and our research takes aim at validating or invalidating some of these. Our first step is to examine prior work that is relevant to a principled analysis of swing dancing.

Prior work can be categorized according to what the various investigators would argue constitutes an understanding of human motion. The spectrum covers: neuroscientific attempts to identify electrical impulses in the nervous system with specific signals in a control theory model [92], phenomenological studies of tradeoffs between speed and accuracy in small reaches [53], kinematic characterization of arm movements [39], optimization approaches which explain simplified full-body movements such as balancing or pedaling [77, 111], and physically-based simulation models for producing animations of human motion [58]. Many of these efforts explicitly consider motion segmentation into primitive behaviors [40, 135], and hidden Markov models are frequently used to segment and classify human movements or intent [86, 108]. Further, investigators have postulated that what makes a movement look skilled, graceful, or natural is either: minimizing the third derivative of position (minimum jerk) [14], minimizing the first derivatives of the joint torques (minimum torque-change) [132], minimizing muscular effort [80], or maximizing use of the extra degrees of freedom available to a redundant manipulator such as an arm [14], or one of a number of other suggested criteria.

Experimental psychology studies perception and control of action by direct manipulation of conditions under which humans are asked to perform specific tasks [53]. In that sense, Chapters 2, 5, and 6 of this dissertation could be classified as experimental psychology.

Detailed classical mechanics models of the human body are available in the text by Tözeren [130]. Sports biomechanics researchers have undertaken to measure and explain solo human capabilities in gymnastics and ballet [84, 85, 103]. Laws' wide-ranging analysis of the mechanics of ballet, for instance, can explain why a dancer's gesture leg in an arabesque turn might oscillate up and down in a troubling way, and how to repair the turn by slowing the rate of rotation and exerting an antiphase lift on the leg [85]. Ballet is a more developed art form than swing dancing, and many of the statements which Laws can prove using mechanics have long been known to experienced ballerinas. Partner dance is a more challenging application of these types of models, because the understanding of human interaction and coordination is in its infancy. For instance, it is not even known how two humans coordinate their actions in an everyday task like lifting a couch together [113].

There is a large and varied literature in the quantitative modeling of human motion and the related task of designing controllers for computer simulations or robots that imitate human motion [12, 50, 77, 78, 94, 133]. Some authors have undertaken to explain the human perceptuomotor system by these efforts [38, 39], while others have restricted their domains to simply producing visually convincing output [58]. We cursorily review this literature here, but quantitative controller models did not play any role in the novel contributions of this thesis. Rather, the difficulty of controlling even a single humanoid model is considerable, and so very few groups have attempted to model interacting humans. Our work instead focuses on qualitative outlines of the means by which leaders communicate with and influence the motion of their followers.

1.2.1 Dynamics and control

Forward dynamical simulations of biological motion, using either simple kinematic or richer physical dynamics descriptions of individuals, have been used extensively in computer graphics for both herds and single actors [18] and these methods are reviewed by van de Panne [133]. The individual dancers are certainly controlled dynamical systems, with measurable inertial properties for a simplified rigid-body description of their segments. The dancers' motion might be described as a closed-loop system with natural $F = ma$ dynamics for the body segments and, for example, joint torque controllers.

1.2.2 Biological dynamics

Our cavalier description in the first sentence of Section 1.2.1 of a human body as actuated by linear joint-torque controllers, something like controllers for torque motors in a robot arm, would not satisfy many biological control researchers. Stiffness control and its generalization, impedance control, have long been favored hypotheses about human motor control systems [59, 94]. Humans can voluntarily generate stiffnesses over a range of two orders of magnitude by co-contracting muscles on opposing sides of a joint [59].

In fact, swing dancers often struggle with learning to use the appropriate stiffnesses. As an example, dance followers are sometimes instructed to have "Barbie arms". Barbie ©Mattel is a classic doll with fairly stiff plastic arms molded with the upper arms at the doll's sides and with the forearms parallel to the floor, held in front of the body. The message is that a follower should not move her arms much relative to her body, even when the leader imposes a force on the arm. Instead of responding to a force on her arm by moving the arm, a follower should transfer that force to her whole body and use the force to turn or travel.

The current view holds that muscles act as tunable springs, in which the motor system selects a length-tension curve [11]. Letting u be an unspecified motor system input, l be the length of a muscle, and f be the tension or the elastic force attributable to spring stretch at that length; a length-tension curve has the form $f = f(l, u)$. The static equilibrium length of the muscle is then $\bar{l}$ such that $f(\bar{l}, u) = 0$. Both external forces like those imposed by a partner, and the inertial and viscous components of force during motion, are disturbances to that notional equilibrium. The equilibrium point hypothesis that the human motor control system actually chooses a final equilibrium point or a sequence of equilibrium way-points, rather that specifying a position or force trajectory, is well-developed if not completely uncontroversial [50]. The equilibrium point hypothesis should also not be viewed as a monolithic entity, for many different types of controllers have been proposed for which parameters of the muscle response are tuned by u, and what controller tunes those parameters in response to what feedback or feedforward model [78, 95, 129].

The impedance view of limb movements actually simplifies in some ways the picture of how two motor systems interact during a dance. Rather than imagining, for instance, that the two partners are each control systems trying to stabilize around some nominal trajectory, or that the leader controls position of their connection points while the follower controls the force connecting them, the two are simply behaving in standard fashion by tuning their muscular springs. The skill that dancers have honed, which we call coordination throughout the manuscript, is prediction of where the total force on their bodies will position them, which is very different from where each one's solo equilibrium would position him (probably on the floor!).

1.2.3 Making complex articulated agents dance *together*

We have seen that there are myriad different control mechanisms that might drive a human-like system, and consensus in the debate over which most faithfully represents real human motor control is unlikely to happen anytime soon [50]. Matarić et al. have compared three controllers, using them to make the Adonis dynamic human torso simulator dance the Macarena [94]. The three controllers considered were a joint-space proportional-derivative

(PD) servo, a Cartesian-space impedance controller, and a joint-space non-linear force field combination of primitives approach suggested by Mussa-Ivaldi's biological models [101]. Each of the methods had serious drawbacks in both the visual fidelity of the resulting motion and in the unintuitive way the Macarena had to be specified. The authors concluded that there is a tradeoff between believability and hand-tuning effort for these approaches [94]. That is, even dynamically simulating solo dance, specified by waypoints and absent any standing balance considerations, is still beyond the capabilities of current dynamic models we have about motor control.

It is certainly not trivial to extend a controller designed for a single person to the problem of interacting people without taking the unrealistic step of applying a centralized controller. The specification of the Macarena dance in Matarić et al. is akin to keyframing, specialized for each of the controllers considered. For partner dance moves, the two generated dancers must maintain contact at their hands, so one can not solve separately for each dancers' position as the hands might not stay connected even if they are touching at every keyframe. Grasping is not permitted, but ideas from non-prehensile, decentralized cooperative robotic manipulation might be helpful [30]. For instance, one might control the leader's hand to compensate for a fixed level of force that the follower will impose [127], or have each individual attempt an equilibrium position just beyond where the other's hand is placed, as in [55]. We are not aware of any existing multi-human simulations of physical interactions which are similar enough to dancing to be adapted to this domain.

1.2.4 Entangled communication

A further complication of the problem of controlling dynamically interacting but independent dancers is the fact that the desired movement is not specified beforehand in lead and follow dance. A restriction that only the leader knows the desired trajectories for both the leader and follower is a *decentralized information structure constraint* [124]. Let each agent i of n agents be a system with control inputs $u_i(t)$ and states $x_i \in X$, with dynamics $\dot{x}_i(t) = f(x_i, u_i)$. If the two or more agents have uncoupled dynamics, then coordination is simply the problem of choosing achievable and non-colliding finite trajectory n-tuples $(x_1, x_2, ..., x_n)(t)$ for $t = [t_0, t_n]$. Call these n-tuples coordinated primitives, or just moves, to borrow a phrase from dance. Frazzoli has described the use of coordinated primitives or moves to facilitate the motion of multiple vehicles [43, 44]. In this framework, transmitted communication of a sequence of moves was permitted. Clearly, transmitting choice of an n-tuple requires much less bandwidth than would transmitting the states continuously.

If transmitted communication is impossible or impermissible, as in covert military operations or human-robot interactions, sensing of the other agents' positions must function as a communication and coordination mechanism. See Fax and Murray for a discussion of transmitted and sensed information flow and its impact on vehicle team coordination [36]. In general, one would not refer to sensing as a communication mechanism. Given a small vocabulary of moves that might be executed, however, sensed variables can be interpreted as communicating the choice of one of those n-tuples. A sensed communication channel could be analyzed in the standard communication framework, where the alphabet is composed of reference trajectory segments; the alternative is to use the sensed information directly as in Fax and Murray's traditional control setting.

Coordination might take place via transmitted or sensed communication. There is a third, most difficult to analyze possibility, that of entangled communication. When agents have dynamics that are coupled in non-negligible ways, the problem of coordination cannot be reduced to the problem of trajectory selection, since there may be multiple control moves that result in a single desired trajectory move. Given the equations $\dot{x_i}(t) = f(x_1, x_2, ..., x_n, u_1, u_2, ..., u_n)$, coordination means that the agents have agreed upon n-tuples of inputs $u_1, ...u_n$, control moves, that result in desirable trajectories $x_1, ..., x_n$, trajectory moves.

When the communication of a type of move can not be partitioned from the control inputs necessary to achieve that move, the term *entangled communication* applies. This means that control and communication objectives are necessarily joined. For instance, if all agents agree to regard a very large u_1 as a signal to choose a certain n-tuple, then that n-tuple must be one where some desirable outcome is achievable given that large u_1. In a sense, widely separated u and x are desirable from the standpoint of minimizing miscommunications, but may be undesirable actual controls and states. Absent a move vocabulary, the problem of decentralized control of systems with interconnected dynamics without transmitted communication has also been addressed in the cooperative robotic manipulation literature [31].

Daniel and McAree suggested that haptic information can be transmitted to humans separably from force. They suggested imposing vibrations above 30 Hz to communicate pure information when necessary, and reserving lower frequency regions for energetic interactions [27]. However, the human motor control system is incapable of creating vibrational messages of this type. In a recent paper, Wagner and Howe attempted to disentangle the signalling and constraint effects of imposing a force on a subject [137]. Section 5.3.2 contains a fuller description of their experiment.

The subjects in the haptic dancing experiments discussed in Chapter 5 and Chapter 6 receive entangled communication messages. Haptics is a relatively immature field, and the two-way nature of haptic interaction makes presenting information to the sense of touch fundamentally different than presenting information to the senses of sight and hearing. Extemporaneous lead and follow dancing is an organic resolution of the conflict between communication and coordination via haptic interaction.

1.2.5 Other work

Dance as a mode of communication has been extensively explored by entomologists studying the waggle dance of various species of honeybee [33]. It is well established that a honeybee's waggle dance communicates the location of a nectar source to follower honeybees, who mimic the dance of the leader. Robotic honeybees have even been constructed which can recruit bees to a location via a waggle dance. Cues are passed to the follower honeybees via sounds, body position as detected visually, vibration of the substrate both bees stand on, and body position detected tactilely. There is substantial physical contact between the leader and the follower honeybees during the waggling run. Through that contact the follower honeybees array themselves behind the leader so that the followers can observe the waggle run angle by placing themselves at that angle [117].

Dancing, like stable walking, has been a standard challenge problem in the robotics domain, although different groups have used the term dancing to describe different sorts of movements. Jalics et al. worked on a solo robot that can dance by stepping to a song's rhythm [64]. Sony's QRIO humanoid robot can dance to techno music or perform a traditional fan dance. We presume the partner dance problem will prove to be more challenging. Some elements of robotic partner dance have been achieved: Setiawan et al. have designed a robot that follows after a human holding its hand [123], and Kazuhiro Kosuge's group at Tohoku University created a ballroom dance follower robot on a wheeled platform [75].

Randomness features prominently in our work on automated choreography. For a dancer's perspective on randomness, consider that Merce Cunningham, a giant of modern dance, has said, "... chance was a way of working which opened up possibilities in dance that I might otherwise have thought impossible" [19]. Cunningham is recognized as a pioneer of choreography created both by random choice and by interaction with computer visualizations of dancers using LifeForms software. "I was working with LifeForms, using chance means to arrange the arms in a way which I'd never seen before. So I tried them in a minimal way and saw that they were possible, so I began to use the arms in a much more complex way" [19]. These are descriptions of unguided chance in creating choreography, which might be more sensible in modern dance than in swing dances. Swing dance still primarily responds to swing music, whereas modern dancers have tried to separate the dance from its natural dependency on the music [141, 140]. Musicality and randomness will be considered together in Chapters 3 and 4.

The individual chapters of the dissertation also contain extensive discussions of the literature related to the specific research questions addressed in each. For a review of computer graphics for animating human motion and previous applications of computing to choreography, see Chapter 3. For a review of work in human-robot interaction and teleoperation, see Chapter 5. For a review of work in the experimental psychology domain of accurate reaching and Fitts' law, see Chapter 6.

1.3 Summary

This overview of partner dancing has focused on the way in which partner dancing is unique among human-human interactions, with the possible exception of martial arts. In partner dancing, there are a number of moves which both partners know, and an even smaller number of positions in which they might start or end moves, and the follower is constantly decoding touch (haptic, tactile, kinesthetic, proprioceptive) cues from the leader. The questions we pose will require insights from several disciplines, and so we introduced some fields in which related work has been pursued.

The contributions of this dissertation, in brief, appear in Figure 1-6. The work presented here will propose and test four hypotheses about partner dance leading and following: that musical meaningfulness is one mechanism by which followers predict moves (Chapter 2), that dance sequences are chosen to express musical meaning in an objective sense such that musical choreography can be created by an algorithm (Chapter 3), that a shared rhythmic vocabulary in random sequences can be used to improve interaction (Chapter 5), and that haptic collaboration enables a two-person team to defeat an individual's tradeoff between speed and accuracy of reaching movements (Chapter 6).

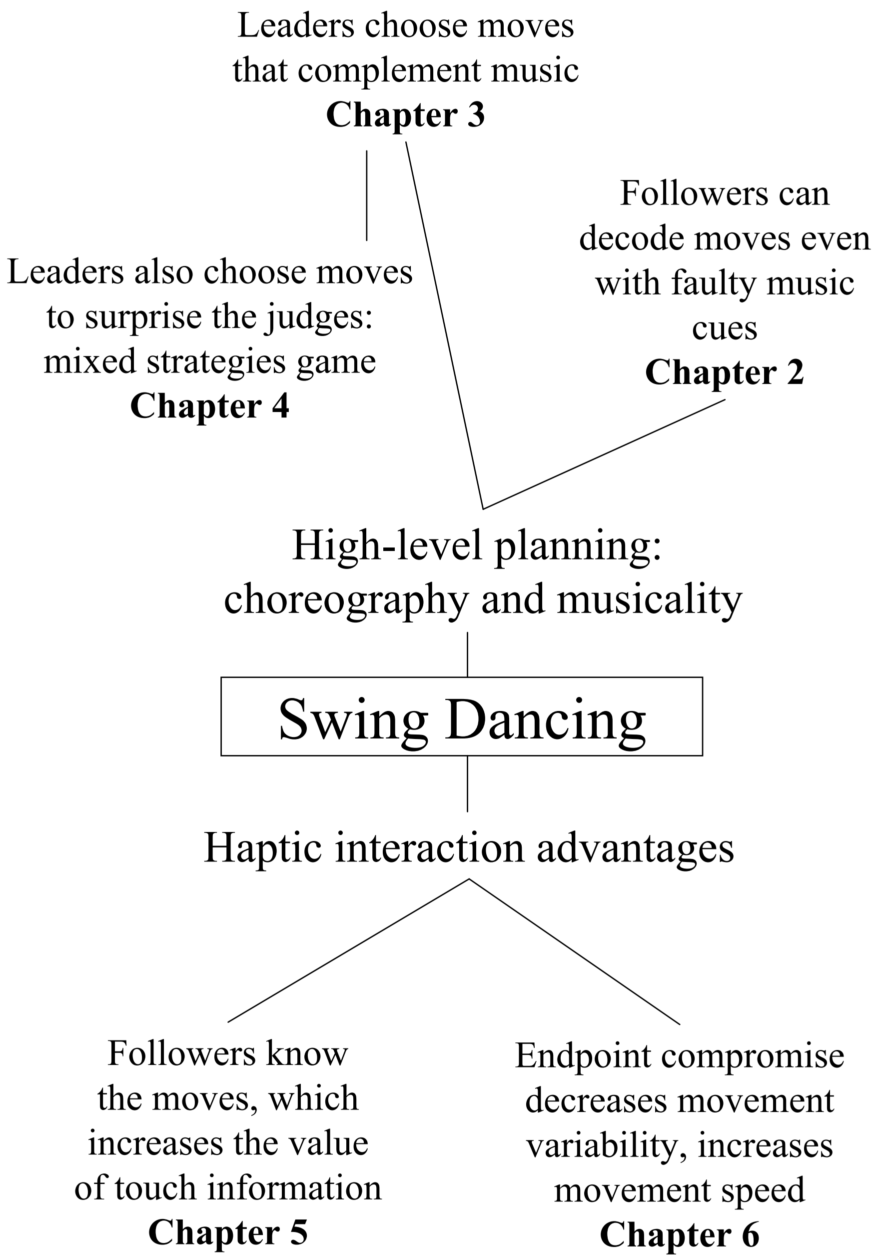

Figure 1-6: Contributions of the dissertation

Chapter 2

Separability of musical interpretation

Sidney Bechet, a famed swing saxophonist, once said, "Jazz musicians have dance in them, and jazz dancers have music in them, or jazz doesn't happen." Any thorough-going analysis of swing dancing must consider the relationship between the music and the dance that accompanies it. This chapter reports an experiment in which the shared musical environment is manipulated to display one song to the leader and a different song to the follower. We began with the hypothesis that obfuscating the cues in musical meaning would impair the communication between the leader and the follower, but instead we had the fortunate experience of discovering an unexpected result. Musical cues do indeed indicate the appropriateness or inappropriateness of specific moves at specific points, but the physical communication between the leader and the follower is robust to differences in song structure. Interviews with leaders and followers who participated in the experiments are key to interpreting the surprising data.

2.1 Introduction

Partnered dance is a successful collaboration mechanism which may point the way to the design and implementation of novel human-robotic interactions [1]. In improvised partnered dance, sometimes referred to as social dance, the leader choreographs the dance in realtime and communicates the chosen moves to a follower. Experiments with a robot leader and human subject followers show that the follower receives the leader's haptic signals and attempts to match her response to his lead from among the moves in her vocabulary (see Chapter 5) [47]. The leader and follower, then, speak the same physical language in which small pushes and pulls on the follower communicate which move the leader wishes to perform.

Does the follower also use information contained in the music to interpret the leader's signals and disambiguate two moves with similar leads but dissimilar musical meaning? It is a platitude of partnered dance to call the music "the third partner", implying that the music cues both leaders and followers to do specific moves. This study investigates that statement.

Clearly, the follower does use music as an absolute rhythm reference. The measure markings on the page of music are discernable as time instants to a listener. In early trials of the experiment discussed in this chapter, some of the music played to the followers had a beat that was slightly out of phase with the music played to the leaders. This rhythm mismatch resulted in a complete breakdown of the dance. With this handicap, the dancers giggled, fumbled, and then stopped dancing, saying, "We just can't do this." Beyond providing a metronome, though, differences in musical expression among bars of music might be essential to disambiguating similar leads. To determine how important this latter effect is to partner dances, we conducted an experiment in contradictory musical cueing.

In Rasmussen's language of human performance models, *skills* function at the lowest level without requiring a person's conscious attention [112]. Keeping time with the musical beats must be at this level of performance because timing occurs subconsciously for dancers and is nearly impossible to consciously alter. *Rules*, however, function as stored programs, released in response to familiar signs in the environment [112]. Rules can be articulated as such by the performer, and indeed, subjects in our experiment can describe how they reacted to musical meaning cues and why. The question at hand is whether the rule-based strategies of the follower for decoding the leader's intent depend critically on the structural and emotional content of the music that both partners hear. Alternative possibilities include: that the follower's rules require only physically communicated signs but not musical meaning, or that the follower does not require any explicit rules but functions entirely at the skills level.

The remainder of the chapter contains: background about musical structure and expression and how dance moves can complement these in Section 2.2, a description of the experiment setup in Section 2.3 and results in Section 2.4, and a discussion of the experiment results in 2.5.

2.2 Background: Musical structure and dance

Lindy Hop is a lead and follow dance composed mostly of a small set of moves that all lindy dancers recognize. Some of these moves are: swingouts, tuck turns, charlestons, and sugar pushes. Different moves suit different expressions in the music; for instance, charlestons are appropriate for very staccato or exciting sections, while sugar pushes and swingouts suit more mellow sections [musicaldancing.mpg]. The above are generalizations, but they provide the basis for expecting that the communication between the leader and the follower may depend in part upon their shared understanding of the music they are dancing to. To be clear, when we refer to communication between the partners, we are considering the case of an improvised dance. Within a pre-choreographed routine, both partners should know all the moves in the dance, and they likely have chosen those moves for musical appropriateness [45].

Lindy Hop is one of the fastest tempo partner dances, often danced at greater than 240 beats per minute, or 4 beats per second. At this tempo, the follower may be receiving and reacting to semantic instructions from her leader every 0.5 seconds or so. Anticipating dance moves based upon the emotional and structural content of the music, if this is possible, may be presumed to have a high value.

Dancers get various types of information from the musical accompaniment as they dance. The tempo and the placement of the beats in time determine the tempo and placement of the dancers' footfalls. This information is so fundamental that if two swing dancers hear music differing in tempo or shifted by anything other than an even integer number of beats they will find it close to impossible to coordinate their movements. The eight-beat bar placement communicates a different piece of information. Swing moves are usually eight beats in length and usually the note at the top of an eight-beat bar, called the 1, is emphasized musically and indicates a time to begin a basic move.

Across larger phrases composed of eight-beat bars, the information contained in dance music is emotional, stylistic, or structural. For this experiment, it is only the latter type of coordinating information that will be obfuscated. The tempo and phase of the music, and the location of the 1, will be the same for both dancers while the emotional emphasis will differ. At issue is whether the emotional and structural content of the music help the follower decode the movement signals she receives from the leader.

Swing music usually has a predictable structure. The traditional structure of swing songs is AABA, with each letter corresponding to four eight-count bars or a 32-count *phrase*. While a musician's bar contains four counts, a dancer's bar contains eight counts. Dance bars in Lindy Hop are usually eight counts in length, so an A phrase contains about four moves. The A phrase is played once, then the A phrase is repeated, then a B phrase differing markedly from the A phrase is played, and then the A phrase is reprised.

The outline of a popular swing song, *Opus One*, the Tommy Dorsey band version, is provided in Table 2.1. For instance, the last bar of each four bar phrase is emphasized, and the last two bars of the song are both loud and exciting as would be expected for the dramatic conclusion to a song. Dancers can use this predictability to extemporaneously choreograph their moves while on the dance floor. It might be appropriate to use swingouts and swingout variations during all the A parts of the song and switch to charleston and charleston variations during the B part of the song.

The rhyme scheme, noted in Table 2.1, does not refer to rhyming lyrics, but rather to rhyming musical bars. Within each phrase of a swing song, for instance, there will be an eight-count musical idea, noted **a**, which repeats twice, for the second and third **a**s, and finally the phrase will conclude with a novel eight-count bar, noted **b**. By virtue of its differences from the three repeated **a** bars, the bar **b** will be emphasized in this song.

A blues song will have a different song structure, an arrangement of six eight-count bars or 48-count phrases. Table 2.1 includes for comparison a chart of a jump blues song, *See You Later, Alligator* by Bill Haley and the Comets. The rhyme scheme in blues songs differs from that in swing songs. Within each A, for instance, there will be two different sounding bars, noted **a** and **b**, and then those two bars repeat, for another **a** and **b**, and then the final two bars of the phrase will be different, **c** and **d**. The last two bars, **c** and **d**, of a phrase with blues structure will be emphasized. An example of this in lyric form would be:

> I knew a girl
> who lived up on a hill

I knew a girl
who lived up on a hill
and if I hadn't met her
she would be there still.

Most trained Lindy Hop dancers have this level of explicit knowledge of musical structure, and the few that do not still demonstrate intuitive understanding of these concepts. The last bar of each phrase is always emphasized, and so dancers emphasize those final bars with more exciting moves. In our experiment, different songs with different phrase and rhyme structures will be played to the follower and leader. This experiment is to determine whether followers need the musical structure, at the phrase level, to decipher the physical cues they receive from their leaders.

2.3 Musicality experiment

In order to determine whether musical expression plays a role in the follower's decoding of signals from her leader, dance experiments and interviews were conducted with 5 experienced leader and follower pairs. Some of these were regular partners and some were not.

Hypothesis 1 *Experienced followers use emotional or structural information in the music to interpret signals from leaders and choose which move from a vocabulary is being led. If leaders and followers are hearing different songs, misinterpreted leads should occur and therefore both partners should be able to detect that they are hearing different songs.*

2.3.1 Procedure

Eight songs which differ significantly in musical structure and style but are all common among lindy hop dance tunes were standardized to a 169 beats per minute tempo and 36 phrase length (about 1 minute, 40 seconds) using Adobe Audition 1.0 software. Additionally, the bars were matched so that the 8-beat phrases of each song coincided, aligning the 1 of every song. The songs were mixed, one song to the left channel and one song to the right channel, in a single sound file. Mixing into a single file permitted the songs to be played synchronously. The right channel was played to both channels of a pair of Sennheiser wire-

Table 2.1: *Opus One* versus *See You Later Alligator*, continues in table 2.2

	Opus One (swing)		*Alligator* (blues)	
Bar	Phrase	Rhyme	Phrase	Rhyme
1	A	a	A	a
2		a		b
3		a		a
4		b		b
5	A	a		c
6		a		d
7		a	B	a
8		b		b
9	B	a		a
10		a		b
11		a		c
12		b		d
13	A	a	A	a
14		a		b
15		a		a
16		b		b
17	bridge			c
18				d
19			B	a
20				b
21	A	a		a
22		a		b
23		a		c
24		b		d
25	A	a	solos	a
26		a		b
27		a		a
28		b		b
29	B	a		c
30		a		d
31		a	A	a
32		b		b
33	A	a		a
34		a		b
35		a		c
36		b		d

Table 2.2: *Opus One* versus *See You Later Alligator*, continued from table 2.1

	Opus One (swing)		*Alligator* (blues)	
Bar	Phrase	Rhyme	Phrase	Rhyme
37	A	a	A	a
38		a		b
39		a		a
40		b		b
41	A	a		c
42		a		d
43		a	B	a
44		b		b
45	B	a		a
46		a		b
47		a		c
48		b		d
49	A	a	B	a
50		a		b
51		a		c
52		b		d

less headphones worn by one participant, while the left channel was played to both channels of another pair of wireless headphones worn by the other participant.

Participants, all experienced lead and follow lindy hop dancers, danced, in three paired trials, for a total of six dances. Participants were informed that within each trial, consisting of a pair of dances, one dance would have the same song for both partners and one dance would have a different song played to each partner. Participants were asked to try to detect whether the first dance or the second dance within each pair had conflicting music. If dancers can detect the mixed-song condition at a rate better than chance, that would demonstrate that the lead and follow connection is perceptibly altered as a result of the music mismatch.

An alternative style of experiment would inform the dancers that their partners are listening to different songs and use a combination of video-tape review and interviews to count misinterpreted leads. This design was rejected because the data could be difficult to quantify: among good dancers robustness to small misinterpretations rendered the visual evidence useless. In preliminary video results from these experiments, misinterpreted moves were indeed masked by quick resolution. Good dancers sometimes expect the unexpected, so the interview would likely produce the non-specific response of, "I don't know whether a lead was misinterpreted because I don't know what I meant to lead; I was giving an open-to-interpretation lead."

2.4 Results

Participants were asked to state whether the first dance or the second dance within each trial was the one with conflicting music. If dancers are not using emotional or structural information to anticipate moves, then they would be expected to guess correctly about half of the time, so the test statistics should be distributed as a Binomial(0.5). The results were unexpected in that although followers could detect the mixed-song condition, leaders could not detect the mixed song condition. Accordingly, we revised our hypothesis.

Hypothesis 2 *[Revised] Experienced followers can use emotional or structural information in the music to determine the musical appropriateness of each move is led, but the force and touch cues alone suffice to communicate move choice. If leaders and followers are hearing different songs, the follower can detect that the moves being led seem unmusical but will still follow leads correctly, leaving the leader unaware of the mixed song condition.*

Hypothesis 2 was confirmed at the $p < 0.05$ significance level in a one-sided test [116]. The results shown in Table 2.3 demonstrate that the follower is able to detect the mixed song condition, but the leader is not able detect the mixed song condition. Followers correctly detected 13 out of 16 mixed song conditions ($p = 0.0106$). Leaders correctly detected 8 out of 16 mixed song conditions, so their responses do not differ from the null expectation that half of the song conditions will be correctly detected.

Although the participants were trying to perform well at this task of detecting which of the two dances in each trial had the mixed song condition, each individual dancer ultimately did not receive a valid estimate of whether he or she could detect that mixed condition. Because each dancer danced only three songs, the resulting statistical test did not have the power to determine whether individuals could or could not detect the mixed song condition.

2.5 Discussion

It is reasonable that followers should be better able than leaders to detect the mixed song condition. In interviews, all of the followers expressed that following accurately was of more value to them than dancing in a musically appropriate way. When these values collided in

Table 2.3: Detected mixed song condition?

	Follower	Leader
trial 1	yes	no
trial 2	yes	no
trial 3	yes	yes
trial 4	yes	yes
trial 1	no	no
trial 2	yes	yes
trial 3	yes	no
trial 1	yes	yes
trial 2	yes	yes
trial 3	no	no
trial 1	yes	no
trial 2	yes	no
trial 3	yes	yes
trial 1	yes	no
trial 2	no	yes
trial 3	yes	yes

the mixed song condition, each follower tried to hide the clash from her leader by following accurately and sacrificing musicality. This conscious choice being made alerted followers to the mixed song condition.

Leaders could not detect the mixed song condition because followers were able to accurately respond to which move was being led, even in the presence of conflicting musicality. This outcome was unexpected, but is a tribute to the skill of the followers and the leaders in the experiment and the robustness of the physical signalling system with which the leader communicates the selection of moves.

In response to a question about whether she considered her role primarily to be dancing to the music or dancing with her leader, one follower had this to say.

> Follower: I would rather follow the leader because he's the one that's in charge. I try to make it work with the music as best I can, but, [I follow the leader even when his lead seemingly conflicts with the music] because you don't want to be all over the place if he's like, "I'm trying to do, I'm trying to work."
>
> Leader: He's already off. She's just making it harder for him [if she tries to match the music when his moves do not]!

The leaders were not often bothered by the conflicting music during the conflicting music dances. In general, each leader felt that his follower agreed with him in the interpretation of the music, even when the music actually differed between them.

> Leader: I danced the heck out of those songs. It's like if you go back and match it up, I'm like, yeah, I'm rocking this song!

However, one leader did notice that his follower was dancing throughout one dance with a different feeling than he was getting from his song. Here, the leader is remarking about the way in which the follower danced the correct moves rather than about which moves she interpreted from his lead. [interview.mp3]

> Leader: Like, the song was kind of Charleston-y, but she wasn't feeling Charleston-y, so I kind of I figured they were different, because it was a different feel.

In one notable trial, both partners agreed easily that a particular dance was to the same song. Within that dance, the follower hijacked a move, taking the leader's role for a moment to choose the next move. Follower hijacks are relatively rare in Lindy Hop. A follower will usually hijack only when there is an extremely compelling musical or emotional point she wants to make. There were two hijacks observed within the 32 dances of the experiment. One of these hijacks is shown in an accompanying movie file [hijack.mpg]. Because she hijacked, the leader had the opportunity to receive information about how she was interpreting the emotional and structural content of the song. In one case, her ideas matched what he was hearing, so he concluded rightly that they were dancing to the same song.

> Leader: This song, I think it was just because of where she hit a break, so I figured, there was a break at the same point.
>
> Interviewer: Did you lead a break?
>
> Leader: No.
>
> Follower: She just did it.
>
> Leader: She just did it, but it was in my music. And so like, she did something separate and there was like something going on in the music, so I figured it was the same song.

2.6 Conclusion

This experiment demotes the importance of the musical content of the environment in enabling coordination between leader and follower. Shared environmental cues may in some cases be unnecessary for coordination of agent teams. However, the musical environments presented to the leader and follower only differed in the highest level information; at the timing and the bar structure level all the music was synchronized. This implies that environmental cues are of more importance in coordinating human actions at the subconscious *skills* level, where timing changes clearly disrupted coordination, than at the *rules* level, where conflicting cues were successfully handled by the followers [112]. Further, this experiment demonstrates that music can clearly encourage or discourage certain moves, because followers were able to detect when the moves they were being led into did not correspond to the music. Rhythm and tone are underutilized communication modes that will prove useful for human teams, human-machine interaction, and even teams of autonomous agents.

Chapter 3

An automated choreographer

Musical interpretation is one of the primary axes on which Lindy Hop dancers judge their performances. Expert Lindy Hop dancers frequently analyze and dissect their musical motivation for choosing certain moves or for varying their footwork patterns. We leverage the aesthetic judgments of humans to imbue a computer with the same sensibility. In particular, musicality, the impression of similarity between bars of music and the chosen dance moves, can be translated from an aesthetic judgment to numerical concreteness. Given a concrete model of musicality in dance, an automated choreographer should be capable of creating dances that demonstrate musicality. Our automated choreographer selects a sequence of swing dance moves, constrained by the allowable transitions of a finite state machine shown in Figure 3-3, to complement a piece of music.

The automated choreographer presented here can choose musically expressive dance moves, with a randomized move selection to preclude prediction by judges or audiences. Dance moves and music bars are scored on emotional attribute scales. At each bar, an optimization model ranks the move alternatives on their musical appropriateness, and then a move is selected randomly from a distribution biased towards the highest ranked moves. This biased distribution is discovered from performer-created dance routines. The optimization model itself can be validated by consistent bias across routines.

Dances that are carefully choreographed and rehearsed actually have greater entropy, or unpredictability, than dances that are improvised with an unknown partner. This calculation appears in Chapter 4. This observation suggests that leaders would prefer to be unpredictable when there is no chance of a miscommunication between leader and follower.

It may be that dancing with an unknown partner requires a slightly more predictable distribution to ensure that followers can predict what the leader is doing. Randomizing the selection of moves while often choosing "good" moves, allows the automated choreographer to demonstrate musicality without being predictable or repetitive. Then, this automated choreographer model will find application in any setting in which optimizing some objective must be balanced with retaining the element of surprise. That the goal of surprising and entertaining dance justifies some random choices will be formalized with a two-person zero-sum game model in Chapter 4.

3.1 Related work

Computers could be incorporated into the art of creating dances in myriad ways, and investigators have been pursuing these multiple lines of inquiry for at least thirty years [105]. To enable a computer to compose a dance, first, the articulated body structure and the desirable qualities of human dance must be represented in computational terms, and second, the computer must produce a dance script which is viewable as an animation, or at the very least is translatable back to human dancers.

On the first point, investigators have often utilized Laban notation, a full-featured dance notation system invented by Rudolph Laban in the first half of the twentieth century [52]. A dance written in Laban notation is analogous to a musical score, even to the level of having dances written in specific movement keys. Laban notation encompasses both the location and structure of movement (what foot is the dancer standing on, which arm is raised to what height, which direction to turn) and the emotive and perhaps intangible qualities of a movement (is the movement meandering or sharply focused, is the movement delicate or powerful) [24]. Other approaches to representing dance have separated the emotional and structural dimensions [118]. For example, Camurri et al. used automated techniques to recognize the emotions of the dancers as they performed the exact same dance with four very different emotions [22].

On the second point, computer graphics techniques for generating all types of human motion have been developing rapidly for the past decade or so. While the earliest attempts at visualizing computed choreography amounted to little more than showing a few widely-spaced keyframes between which dancers were free to interpolate [82], modern techniques are often capable of producing commercial-quality animation [89, 110]. Obviously, the capabilities of any dance composition and visualization system depend largely on the encoding chosen for the dances, so the most sophisticated efforts have addressed these problems taken together.

3.1.1 Motion segmentation and generation

Motion segmentation nearly always underlies the frameworks both for notating and generating motion. The finite state machine we will describe for swing dancing echoes all of: Rose's verb graphs [118], Kovar's motion graphs [76], Kim's movement transition graph [74], Li's motion texton transition matrix [87], ad infinitum. Automated motion segmentation is performed in some of these works, and also considered as a unitary problem in many others [68, 135]. Fod, Matarić, and Jenkins incorporate psychophysical research into the segmentation and representation of movement primitives [40]. Kim et al. have analyzed motion-capture of rhythmic dance, discovering the underlying beat for use in segmentation and calculating Markov transition probabilities between dance moves to generate novel dance sequences [74].

Generation of animated humans from segmented data or segmented models can take one of three basic forms: interpolation, procedural generation, or dynamical simulation [118]. Procedural generation, which encodes rules to drive the positions of the limbs, requires tuning time proportional to the amount of detail needed in the final animation. Dynamical simulation generates motion using controllers and forward simulation of physics-based models of the body [58, 133]. A more complete review of research in dynamical simulation of human motion, including neuroscience and motor control aspects of this work, appears in Chapter 1. Interpolation is the most direct method for reusing motion capture and other animation data, and numerous interpolation techniques exist for reorganizing, blending, overlaying, enhancing, and controlling human motions [76].

Interpolation methods which can be controlled at a high level to visualize arbitrary paths through a state machine are complementary to our work if the methods are compatible with fixed tempo constraints. An automated choreographer which selects motions from a vocabulary to suit a piece of music creates a script, and that script can be made viewable by interpolation. Pullen and Bregler describe a method to control interpolation via keyframing, allowing a user to fill in missing degrees of freedom directly from motion capture [110]. They

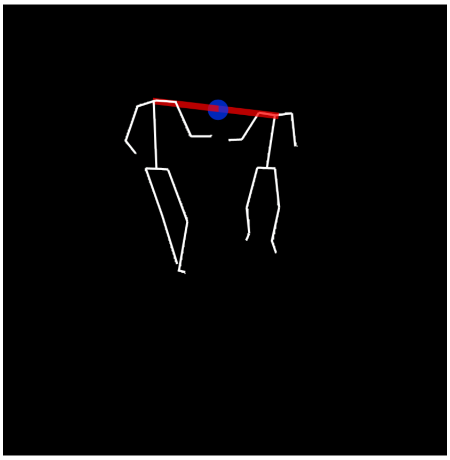

Figure 3-1: End pose of previous move

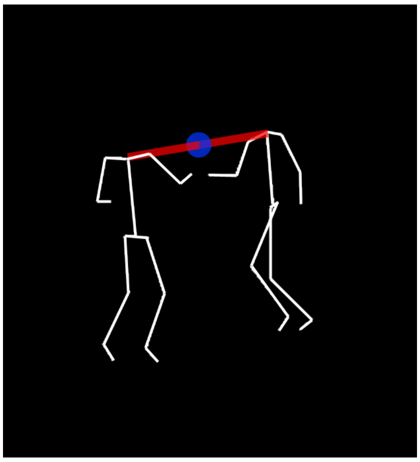

Figure 3-2: Start pose of next move

applied the method to modern dance, keyframing only a few poses of the legs and generating full-body motion with realistic texture. Hsu, Gentry, and Popović used a related method to generate a swing dance follower's motion to complement the captured motion of a swing dance leader [60].

Fortunately, swing dancing is structured such that movement segmentation and the connections in a transition graph are given almost for free. For instance, as discussed in Section 1.1.2, there are only a few well-defined positions in which any move could begin or end. This consistency means we may take fewer pains to blend between moves in a sequence than many other applications might require [76]. Our automated choreographer uses an elementary form of interpolation to display its output. It simply connects motion-captured segments by translating and rotating each move. As pictured in Figure 3.1.1, the translation and rotation is calculated to overlay the center of a line segment (thick line) between the follower's sternum and the leader's sternum, with the centers of the line segments at the same point (large dot). No blending between motions is performed. We have simplified the dance from its most general form, most notably in considering all movements as 8-count rather than the mix of 6-count, 8-count, and other even-count moves that would appear in Lindy Hop.

3.1.2 Computer notation and representation as choreographer-driven tool

Some investigators have used the computer to interactively try out choreography on computer-generated dancers rather than real dancers, as with the software LifeForms [19], or to translate precise dance notation into graphic images as with the software LabanDancer. An excellent review of the long history of these two lines of inquiry can be found in [20]. Methods for the inverse task of translating video or motion capture into dance notation are less developed [21, 145]. In each of these approaches, the computer is ancillary to the actual task of selecting which movements will comprise the dance performance.

3.1.3 Randomness is often a desirable feature

Randomness has long been considered a desirable feature of computerized choreography. However, our approach to balancing musicality with unexpectedness, and in particular our modeling of choreography as a two-player game in Chapter 4 appears to be novel. An early method of computerized choreography was to randomly assemble keyframes in an order, then depend on the wisdom of dancers to interpolate between them [82]. Merce Cunningham used the inverse kinematics solver and random assignments of limb positions within LifeForms software to give him ideas about novel arm gestures [19]. Cunningham also embraced randomness as a means of expanding possibility within performances, for example rolling a die each night to decide the order of pieces for that evening's show. Perlin, within a procedural approach to motion generation, uses pseudorandom noise functions to lend an air of naturalness or personality to CG puppets [109].

Nearly any segmentation and transition graph approach for reusing motion in a novel sequence can be driven in random fashion. Random transitions between frames can be used to create infinitely playing video textures of a candle flame or a waterfall, with appropriate corrections to keep the video texture away from dead ends [122]. Doretto, Chiuso, Wu, and Soatto perform a similar task, infinite novel video from a finite capture, using ARMA models to describe each image within a video texture as a realization of a dynamic process driven by independent and identically distributed noise [32]. Kim et al. produced novel sequences as realizations of a fixed Markov chain, with transition probabilities between dance moves estimated from the captured dance data [74].

Approaches like these could certainly generate novel sequences, but even Merce Cunningham filtered out some of the random gestures that LifeForms suggested. That is, a measure of goodness or appropriateness should also be incorporated into automated choreography. The contribution of this chapter is to incorporate both the desirable excitement of random selections, and a sense of musicality in that dance moves harmonize with the soundtrack, into an automated choreographer. In the following chapter the interaction between the dancers and their audience or judges is modeled as a game, the solution of which clarifies why randomness in choreography should be valued.

3.1.4 Stylistic and emotional warping

Dance is often invented with the goal of communicating an emotional or stylistic impression to the audience. Automated techniques for interpreting the emotional content of dance have shown some success as judged by concurrence with spectator interpretations [22]. EyesWeb is a system for interpreting the dance content of 2D video data along some custom-designed style axes, a few of which correspond to Laban parameters [21]. The stated goal of EyesWeb is to provide live feeds to enable dancers to drive interactive music and staging enhancements. An installation of this type, like the interactive dance club at SIGGRAPH [131], creates an automated musician guided by the dance instead of an automated dancer or choreographer guided by the music.

Consider now producing animation to reflect emotional or stylistic content. Given a neutral motion captured movement, some groups have attempted to enhance the motion with emotive accents [6]. Emotional accents directly correlated to Laban notions of effort and flow have also been added to keyframed movement [24]. Style machines go beyond fixed descriptions of moves and styles to identify from a collection of motion captured data both the recurring choreographic elements and the most salient style parameters [17]. The resulting style machine can be driven with noise, scripts, or video to create new sequences or to convert novice ballet motion to that of an expert.

3.1.5 Computerized choreography

It is perhaps surprising that in none of the work described above was the musical accompaniment to a dance considered at all, except in Kim et al. which made limited use of rhythm [74]. Instead, works which bear the most similarity to our designs for an automated choreographer are earlier automatic visualizations intended to complement a music soundtrack. Generally, efforts of this type have either analyzed waveform properties of audio signals, or interpreted the literal score which a MIDI file encodes. An elementary example of the former type are the Bars and Waves visualizations available in Windows Media Player, which display the frequency spectrum of the song or mimic an oscilloscope [26]. Lytle's Animusic is a particularly well-developed example of the latter, in which MIDI files are visualized with rich animations of musicians playing each note [88, 89]. Wang et al. controlled a cube of springs according to some MIDI analysis [139]. In the latter category also are interactive authoring tools such as MIDIArt, which allow users some freedom to alter the visualizations (here, colored circles and embellishments) of MIDI notes [42]. Cardle et al. combined MIDI and audio analysis tools to transform a keyframed or motion-captured sketch into a music-driven final animation [23]. The latter work did not consider the high-level planning problem of choosing which movements would complement the soundtrack, as does our automated choreographer.

3.2 Introduction

Choreography is the art of arranging dance moves, usually with the objective of creating a dance that reflects the music danced to and communicating expressively to the audience. An automated choreographer might allow hands-off computer animation of large group dance scenes; or just provide musically meaningful movements for a simple screen display to accompany the playing of music files. We introduce a model for automated choreography built from first principles and actual choreography data.

The choreography data available might be the result of an optimization process. The expert dancers might choose from a set of moves the best moves to complement the music, given the constraints that one move must end where the next begins. Then inverse optimization, which aims to identify optimization parameters from given optimal decisions, could be applied to determine the leader's objective function precisely [4, 35]. Given the leader's objective, the audience could predict exactly the sequence of moves that a perfectly optimizing leader would choose for a certain song.

A choreographer, though, should have many musically appropriate outcomes rather than a single optimal one. An automated choreographer must have a random element, so that it can produce different dances for the same song, as a real dancing couple would. Our model uses optimization to rank the possible moves at each bar, and then biases the randomized move choice toward highly-ranked move transitions. The automated choreographer can be called a rank-biased randomizer for choosing the next move. The bias, the distribution over highly ranked moves, is identified from data. The optimization process that produces the ranks is based on domain knowledge, and verified by identifying reasonable and consistent rank distributions from independent datasets.

In fact, what we choose to model as a random element could be the result of inputs we have not modeled, such as the difficulty of particular moves or the difficulty of particular transitions between moves. The data in this chapter come from choreographed routines, in which the dancers have had ample time to consider what moves are musically appropriate. The influence of musical expressiveness on choreography should be detectable, so that moves which are particularly appropriate for certain sections of music are more likely to be chosen than less fitting moves. Recall that Chapter 2 confirmed that among expert dancers, musical expressiveness is a real phenomenon. In order for followers to experience and detect the dissonance between their music and the leader's choreography, leaders would have to have been leading musically meaningful moves [46].

Ultimately, an algorithm for choreographing musically expressive dance must turn to artistic judgement for validation. Lindy Hop has a restricted set of named movements that the leader sequences to create choreography. A move transcription of a 2001 American Lindy Hop Championships (ALHC) routine to *Frenesi* is used here as both: training data to fit a probability distribution, and validation data for the a priori optimization model. A move transcription of a 2003 routine to *Yes, My Darling Daughter* for the same contest validates the rank probabilities learned from the ALHC training set.

The remainder of the chapter is organized as follows: a description of the automated choreographer in Section 3.3, data fitting and model validation in Section 3.4, and discussion and applications in Section 3.5.

3.3 Optimize, then randomize

To create a ranking of the musical expressiveness of moves, an optimization model ranks each of the move choices separately for each bar of music. To match movement to musical expression, attributes of both the movements and the music are evaluated on numerical scales. We have chosen the attributes *loud, exciting, staccato, emphasized* and assigned numerical values between 1 and 10 to each eight-count bar of music and, independently, each eight-count piece of choreography. The assigned scores for the song *Frenesi* are in Table 3.1. The attribute *emphasized* refers to a break in the music; a 10 on the scale of emphasis might be a very loud drum hit followed by a few beats of silence. Dance moves high on the emphasis scale might involve a dramatic pose held for a few beats. The musicality parameters for each song and dance are fixed. The transcribed dance data are used to estimate a probability distribution over the most musical moves as ranked by an optimization procedure.

The optimization move ranking model restricts consideration to a set of 42 Lindy Hop moves, which are listed in Table 3.2 with each move's assigned musicality scores.

3.3.1 Optimization model to rank moves

Using the weights w_j for the attributes indexed by j, let the distance or dissimilarity between a move with attribute scores m_j and and a bar of music with attribute scores s_j be:

$$d(m, s) = \sum_j w_j * (m_j - s_j)^2 \tag{3.1}$$

The squared distance was chosen to penalize large deviations in one of the dimensions, say, just in the *staccato* dimension, more heavily than small deviations in many dimensions. Let a dance M, for a song S with N bars, be a sequence of N moves. Let the attribute scores for the nth move and nth bar of music be s_{nj} and m_{nj}. The total dissimilarity between the dance M and the song S is:

$$d(M, S) = \sum_j w_j * \sum_{i=1}^{N} (m_{ij} - s_{ij})^2 \tag{3.2}$$

Lower and better scores on this metric are assigned to dances which have attribute scores similar to the song's attribute scores, subject to the attribute weighting.

The optimal dance for a song S, might be defined as the feasible dance M minimizing the metric $d(M, S)$. Only some sequences of moves M are feasible. For instance, a lindy circle finishes in closed position, with the dancers standing very close and the leader's right arm wrapped around the follower. A whip, which starts from open position in which the dancers are at arm's length and facing each other, could not follow a lindy circle. See Figures 1-1 and 1-2 for the two main poses in which moves may start and end: closed position and open position. See Figure 3-3 for a state machine linking the 42 moves used in our model.

Minimizing $d(M, S)$ subject to move continuity restrictions can be posed as a shortest

Table 3.1: Song attribute scores (abbreviated)

Song: *Frenesi*	*Loud*	*Ex.*	*Stc.*	*Emph.*
bar 1	4	3	4	3
bar 2	4	3	4	3
bar 3	4	3	4	3
bar 4	4	5	5	7
bar 5	4	3	4	3
bar 6	4	3	4	3
bar 7	4	5	4	3
bar 8	4	6	5	7
bar 9	6	7	7	8
bar 10	6	7	7	8
bar 11	6	7	7	4
bar 12	6	8	8	9
bar 13	4	3	4	3
bar 14	4	3	4	3
bar 15	4	3	4	3
bar 16	4	5	3	5
bar 17	2	2	2	1
bar 18	2	2	2	1
bar 19	2	2	2	1
bar 20	2	2	2	8
bar 21	9	9	10	5
bar 22	9	9	10	5
bar 23	9	9	10	5
bar 24	9	10	10	7
bar 25	4	4	3	3
bar 26	4	6	6	4
bar 27	4	4	3	3
bar 28	4	6	6	4
⋮				
bar 49	7	5	4	3
bar 50	7	5	4	3
bar 51	7	5	4	3
bar 52	7	8	5	8
bar 53	5	5	5	3
bar 54	5	5	5	3
bar 55	5	5	6	3
bar 56	4	4	8	8

Table 3.2: Move attribute scores (abbreviated)

Lindy Hop move	*Loud*	*Ex.*	*Stc.*	*Emph.*
foxtrot	3	3	1	1
circle	5	7	2	4
freeze	5	10	5	10
swingout	5	8	6	4
closedfootwork	5	8	7	7
popturntoclosed	4	4	3	3
20sCharleston	6	8	9	7
Balboa	1	1	1	2
crosshandcircle	4	3	3	5
charleston	7	8	8	7
flashkicksjockey	7	5	7	8
eastcoast	4	3	3	3
promenade	3	4	6	2
sendout	5	6	5	4
popturn	4	5	3	2
tuckturn	5	3	6	6
whip	5	5	3	4
whipleaderspin	5	7	4	5
whipinsideturn	4	4	4	4
whipoutsideturn	6	6	4	7
underarmturn	3	3	3	2
hegoesshegoes	5	7	8	2
sugarpush	4	4	7	7
rightsidepass	3	3	3	4
⋮				
pizzatoss	8	8	6	8
texastommy	5	5	4	8
crosshandwhip	4	7	5	4
judoflip	8	10	8	10
knickerbocker	8	10	6	10
getbackhere	6	5	3	4
promenadeopen	6	5	5	4
shortygeorge	4	5	6	5
tabbythecat	6	6	4	8
sugarpushfreeze	5	7	6	8
switches	5	5	4	3
circleduckunder	6	8	6	9
leadersplits	5	6	6	6

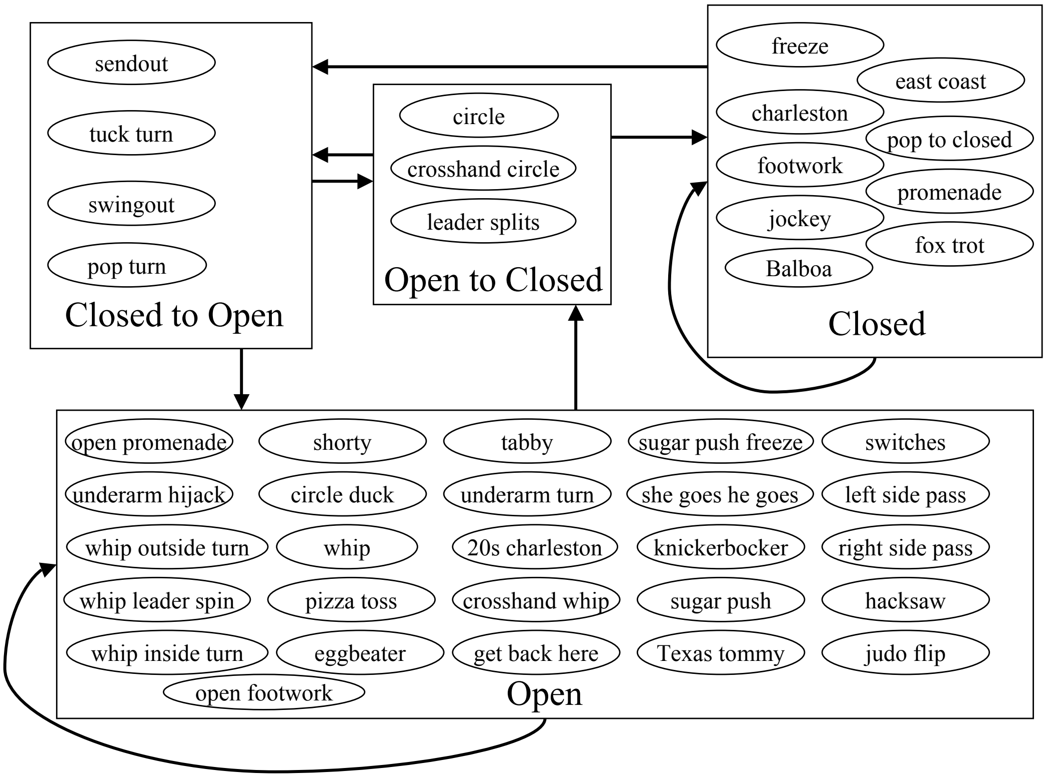

Figure 3-3: Dance choreography is a finite state machine. Reciprocal links exist between all moves in each self-linked box.

path problem. Let flow at a node $[nk]$ represent the choice to use dance move k from the dancers' vocabulary during musical bar n. Allowable transitions are connected with arcs, so that if move a ends in the state that move b starts from then the arc connecting node $[na]$ with node $[(n+1)b]$ will be present for all $n < N$. The node costs are the one-step distance metric d as in (3.1), while the arc costs are all zero.

At each bar of music, a shortest path problem starting from the current dance move must be solved. These small linear programs can be solved quickly and reliably using AMPL/CPLEX [41]. In Section 3.3.3, we will consider randomizing the choice of move rather than selecting the very shortest path.

3.3.2 Horizon h repeated shortest paths method

The optimization horizon in the repeated linear programs solved above is the remaining length of the song, from $n...N$. The state machine is highly connected, because there is at most one intermediate move separating any start move from any desired target move. In this section we consider whether a shorter optimization horizon would suffice. If, for example, there were a direct link between every pair of moves in swing dance, then optimization horizon $h = 1$ would suffice, because each move would impose no restrictions on the possible moves to follow.

The shortest path problem arranged in stages could be solved by a sequence of horizon h shortest path problems. Specifically, a horizon h shortest path problem derived from the optimal dance setup would insert a dummy node after h bars of music, and solve the resulting shortest path problem. Then, after committing to the first dance move on the shortest path with horizon h, a new horizon h shortest path problem would be created from music bar 2 to music bar $h + 1$. This is illustrated in Figure 3-4.

In experimenting with shorter planning horizons, we have found that using a horizon of $h = 5$ is sufficient to find the optimal choreography, while horizons of $h = 4$ or less sometimes result in a suboptimal choreography path. However, this result is only a quirk of the particular distance metric and set of dance moves chosen, as will be shown. Figure 3-5 shows the ratio of the cost of a dance planned with a shortened optimization horizon to the minimum cost found by planning with horizon N. The cost does not always decrease monotonically with increased horizon, and the penalty for shorter horizons is not larger than 12% in these examples. No randomization step was used, so the scores shown are

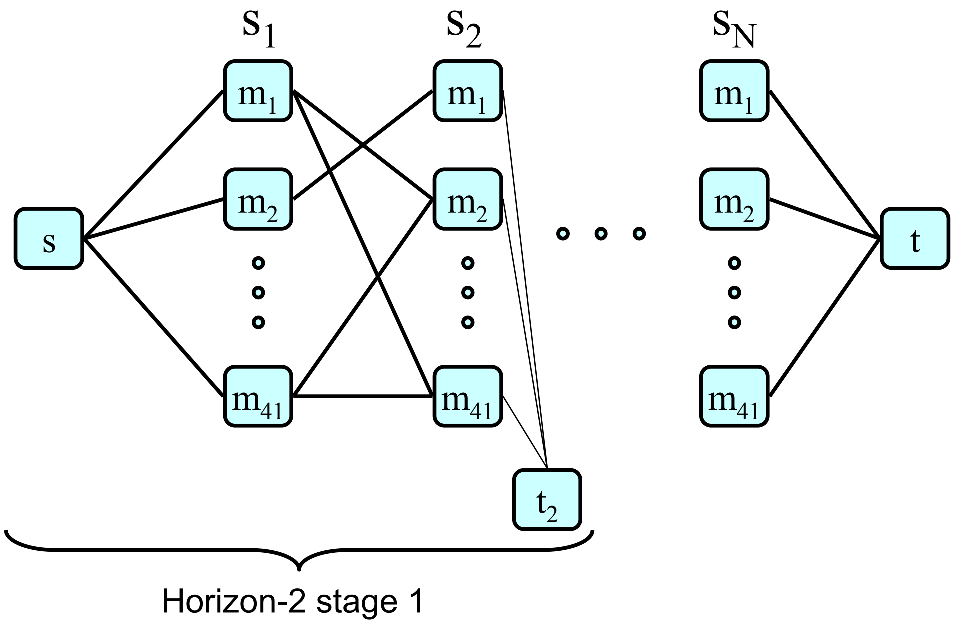

Figure 3-4: Choreographer as sequence of horizon h shortest path stages. The move selected at each bar of music s_n is that chosen by a shortest path solution up to bar s_{n-1+h}.

those of the horizon h shortest path.

Davis and Impagliazzo have applied competitive analysis techniques from the online algorithms literature to prove bounds on the approximation ratio achievable by various classes of algorithms for graph problems [15, 29]. A fixed priority algorithm is one that, given the input instance, chooses a fixed order in which to examine the data items and make an irrevocable decision for that item. For instance, an edge might have to be irrevocably included or not included on the short path to be reported by the fixed priority approximation algorithm. An adaptive priority algorithm may rearrange the order in which to examine data items depending on what it has seen. Davis and Impagliazzo show that no fixed priority algorithm can even approximate the solution to a shortest path algorithm. Djikstra's algorithm for shortest path is an adaptive priority algorithm.

The horizon h repeated shortest paths method fits the definition of a greedy algorithm given in [16]. The horizon h repeated shortest paths method is nearly a fixed priority algorithm for the overall shortest path, except that an additional fixed set of data items H_h is available before the irrevocable decision is made on each edge. Let D_n be the set of data items pertaining to music bar n, that is, a list of the edges from moves at music bar $n-1$

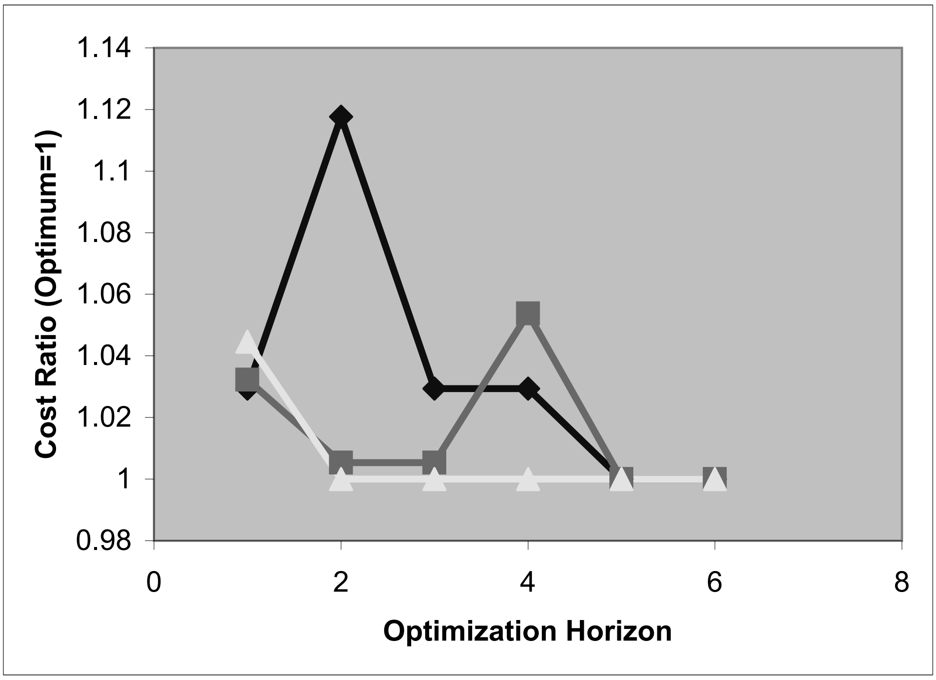

Figure 3-5: Shorter optimization horizons result in higher total dance cost. The ratio of horizon h cost to minimum cost is shown, for three sample choreography problems.

and their weights $d(m, s_n)$ from equation 3.1. Then $H_h = \{D_n, D_{n+1}, \ldots, D_{n-1+h}\}$ is the data available at the time the irrevocable decision about which move to choose for music bar n is made. By counterexample, we demonstrate that even the special structure of the choreography problem in which any two moves are separated by at most one additional move, does not enable shortening the optimization horizon in the general case. This does not rule out the use of an estimated cost-to-go function, as in approximate dynamic programming [10].

Theorem: The horizon h repeated shortest paths method for the general choreography problem can not be guaranteed to find the optimal choreography path, for any non-trivial h. A non-trivial h is an h less than the full length of the song.

Proof: We demonstrate a problem instance in which the Solver will have committed to a suboptimal path, with a total number of stages N that is 1 greater than the optimization horizon h. Consider a simple example, in which there is only one move of each of the following types: Begin Closed-End Closed (BCEC), Begin Closed-End Open (BCEO), Begin Open-End Closed (BOEC), Begin Open-End Open (BOEO). The instance is drawn in Figure 3-6 for any horizon size greater than 2. The solid lines each have weight 0, the

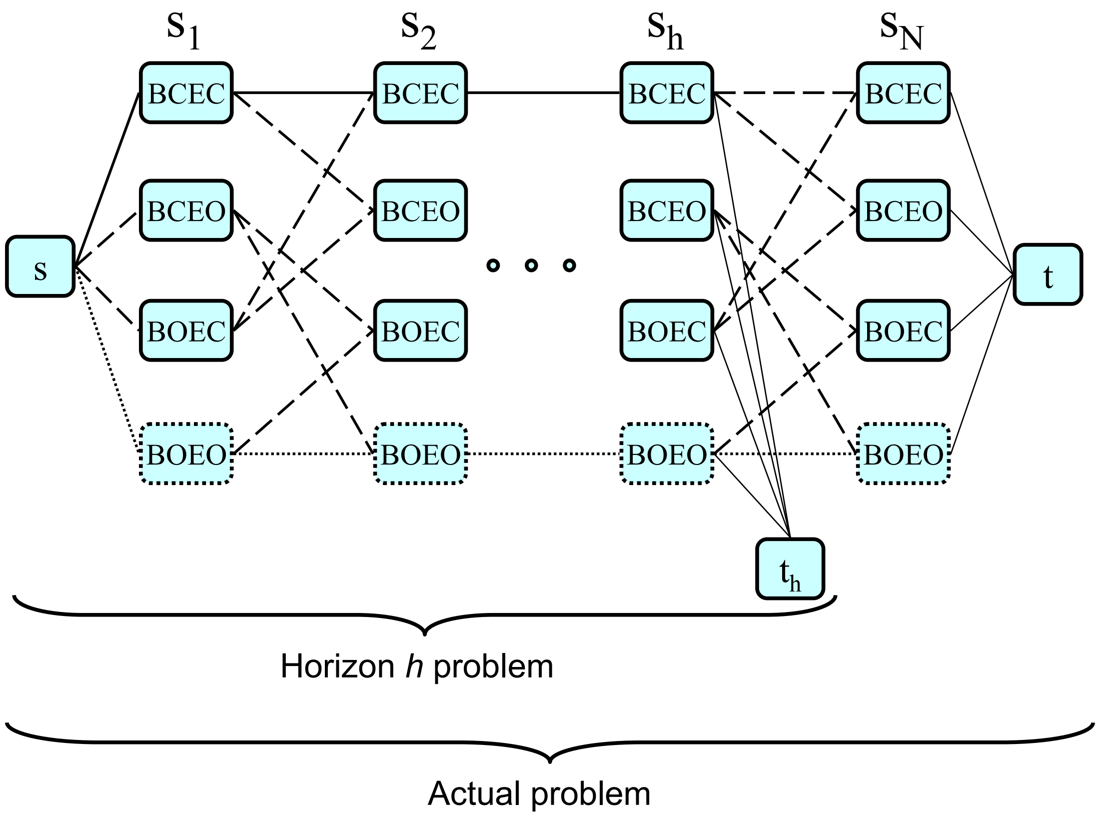

Figure 3-6: Problem instance in which a horizon h shortest paths method will commit to a suboptimal path. The weight on each solid line edge is 0; the weight on each dotted line edge is 1; and the weight on each dashed line edge is $3h$. All edges are directional from left to right.

dotted lines each have weight 1, and the dashed lines each have weight $3h$. The solution to the horizon h shortest path problem ending at t_h is to follow the path across the top of the diagram through the BCEC node at every bar of music, with path length zero. The true shortest path is the path marked by dotted lines across the bottom, through the BOEO node at every bar of music. The generalization of this counter-example to a horizon length which is any fraction of the song length N is trivial.

3.3.3 Rank-biased randomization

In order to create surprising dances, the automated choreographer will randomly choose a move at each bar of music. At the first bar, the optimization model will rank all the available move choices. None of the moves are eliminated from consideration by the connectivity constraints because the dance may start either in closed or open position. Call the resulting list of ranked moves r_1, and let a probability distribution $\{r_1(i), p(i)\}$ with $p(i) \geq 0$ and $\sum_i p(i) = 1$ be given. The distribution presumably favors highly-ranked moves, which are the most musically appropriate. Later we will identify this distribution for Lindy Hop from data.

The distribution $p(i)$ is given over the ranks of moves rather than over the moves directly. A new rank-ordered list of moves r_2 will be calculated for the alternatives at the second bar of music. The automated choreographer will again select from the distribution $p(i)$ over the new list r_2.

Choreographing a dance to a song, then, requires sequential randomizing over the ranked alternatives to a linear program with some stochastic data. The stochastic datum at each stage is the move chosen for the previous bar of music. An accurate ranking of moves would require solving a multistage stochastic linear program (MSSLP) [69]. We take the shortcut of solving deterministic linear programs at each nth bar of music. Given the move chosen at bar $(n-1)$, we assume that the single best move will be chosen for the remaining bars of the song. Since the connectivity of the state machine in Figure 3-3 is high, committing to any particular move at any stage does not heavily restrict choices later than the immediate successor.

At each stage n, the automated choreographer chooses a move in the rank-ordered list r_n from a discrete distribution $\{(r_n(i), p(i))\}$. See Figure 3-7 for a graphical outline of the automated choreographer algorithm. For instance, if at the first bar the randomized move

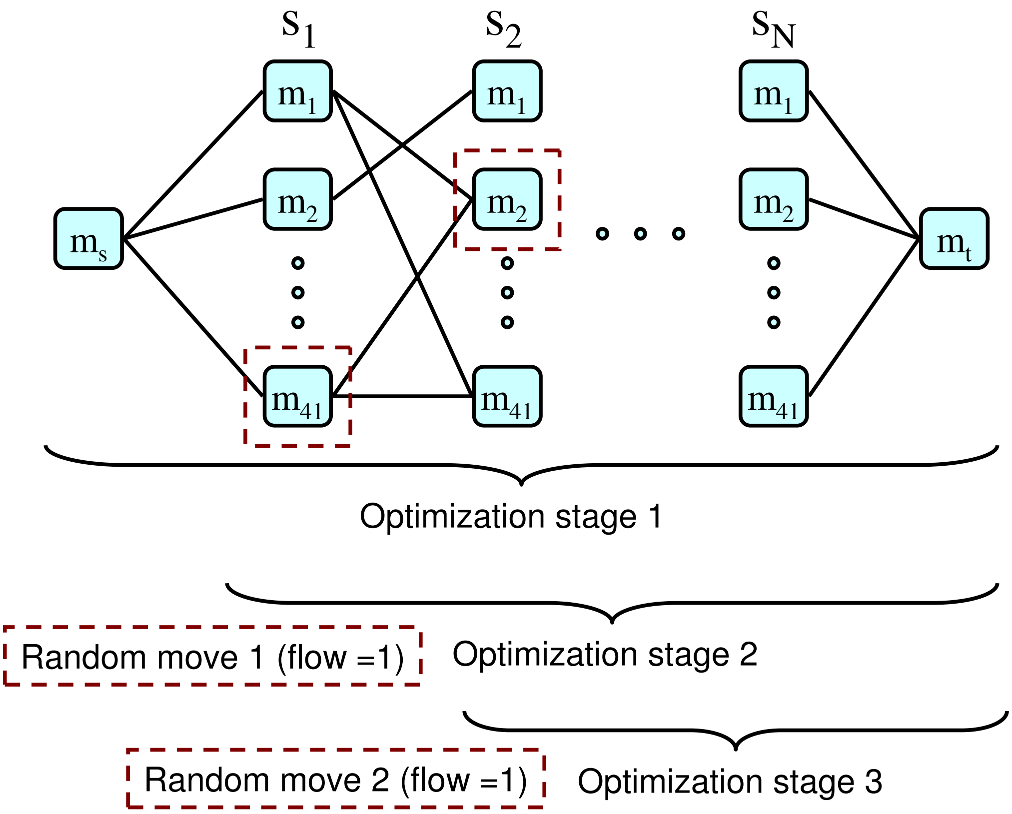

Figure 3-7: Choreographer as sequence of shortest path move ranking stages. The randomized move selection at each bar of music s_n is biased by the ranks assigned by optimization.

chosen is m_{41} = *circle*, then the next step is to minimize (3.2), a shortest path problem, from bar 2 to bar N. The move chosen at bar 2 must be a feasible successor move to a *circle*. This constraint is implemented by setting node [1,41] as the source for the next shortest path problem minimizing (3.2), where the sum is taken from $i = 2$.

The move that best fits bar 2, say, m_8 = *swingout*, is then removed from the problem by eliminating node [2,8]. (The choice shown in Figure 3-7 at bar 2 is different.) Solving again the shortest path problem minimizing (3.2) from bar 2 to bar N identifies the second best successor move, say, m_6 = *closed footwork*. Ranking all move alternatives for bar 2 is called optimization stage 2 in Figure 3-7. Each optimization stage n repeats these steps: solve a shortest path problem and then remove the node chosen as best at bar n; until the shortest path problem is infeasible. In implementation, the automated choreographer first selects the rank i of the move to be used for bar 2, and then the iteration described in this paragraph is halted after i steps, that is, after the move ranked ith has been identified.

3.4 Data fitting and model validation

The shortest path model described in Section 3.3 suggests a method for ranking moves from an actual dance routine according to the musical aptness of the moves in the routine. Transcriptions from video of two routines from the ALHC, one to *Frenesi* and one to *Yes, My Darling Daughter*, were taken, and from these the ranking of the actual moves danced was calculated. From each actual move in the transcribed dance data, say, a *sendout* in bar 2, all possible next moves were ranked. Say the actual move danced in bar 3, a *whip*, is ranked fourth among alternatives. Then the sequence of the ranks of the observed moves comprises a new dataset, the routine's ranks. If the musicality optimization model has merit, then the ranks of the observed moves should be biased towards the highly-ranked alternatives. In Chapter 4, we consider a uniform distribution over a set of highly-ranked alternatives, but then show that these different models are not mutually contradictory.

The automated choreographer requires only a few parameters to be fit from the data; these are the observed probability distribution over the move rankings. How frequently is the very best next move chosen, and how frequently is the second-best next move chosen? This probability distribution will be used to propagate move choices in the automated choreographer.

3.4.1 Results: optimization model validation

The ranks of the moves in the *Frenesi* routine were calculated, using equal weights $w_j = 1$ for $j = 1, 2, 3, 4$ in distance metric (3.2). For comparison, a random feasible dance unrelated to the musical structure of the song *Frenesi* was compiled from the same set of 42 moves, and the ranks of the moves chosen in that random dance were calculated. Table 3.3 demonstrates that the *Frenesi* routine chooses transitions that rank high in musicality, while relatively few moves in the random dance rank high in musicality. Table 3.3 also shows that unmusical moves very distant from (dissimilar to) the music bars are rare in the actual dance and occurred frequently in the random dance.

Table 3.3: Musical moves are highly ranked

	Frenesi dance	Random dance
Among best 5	30 of 56	13 of 56
Among best 10	42 of 56	22 of 56
Among best 15	49 of 56	30 of 56
Distance > 50	2 of 56	16 of 56

3.4.2 Results: bias towards highly ranked moves is consistent across routines

A second move and song transcription from the same contest were used to rank the moves chosen by a different pair of dancers, using the musicality optimization model and weights of the previous example. This test data routine was danced to the song *Yes, My Darling Daughter* as performed by the Eddie Reed Big Band. The distribution of musicality ranks of the dancers' chosen moves was compared to the calculated rank distribution from *Frenesi*, and both were also compared to the distribution of musicality ranks of a random dance.

The nonparametric Kolmogorov-Smirnov two-sample test was used to determine how likely or unlikely would be each pair of observed datasets if they came from any identical underlying distribution. The Kolmogorov-Smirnov test is formulated for continuous distributions. However, the discrete move ranks as calculated by the optimization model are ordered over the range from 1 to 30, so they are similar to continuous values known to very low precision. The K-S test is conservative in the case of discrete data, so the actual p-values are somewhat lower than those calculated [25].

Table 3.4.2 shows that the null hypothesis, that the *Frenesi* dance routine's ranks and the *Darling* dance routine's ranks are drawn from the same distribution, can not be rejected ($p = .4144$). The random dance routine ranks are extremely unlikely to have come from the same distribution as either the *Frenesi* ranks or the *Darling* ranks ($p = 0.00003918$, $p = 0.0042$). This demonstrates similarity between the rank biases of the *Frenesi* dancers and the *Darling* dancers that can not be an artifact of the optimization model's ranking. That is, the bias toward highly ranked moves is consistent across the routines, but the random dance was likely not drawn from the same distribution over move ranks.

Table 3.4: Two-sample K-S test with null hypothesis that both datasets come from the same distribution

	Darling ranks	Random ranks
Frenesi ranks	**0.4144**	0.00003918
Darling ranks	.	0.0042

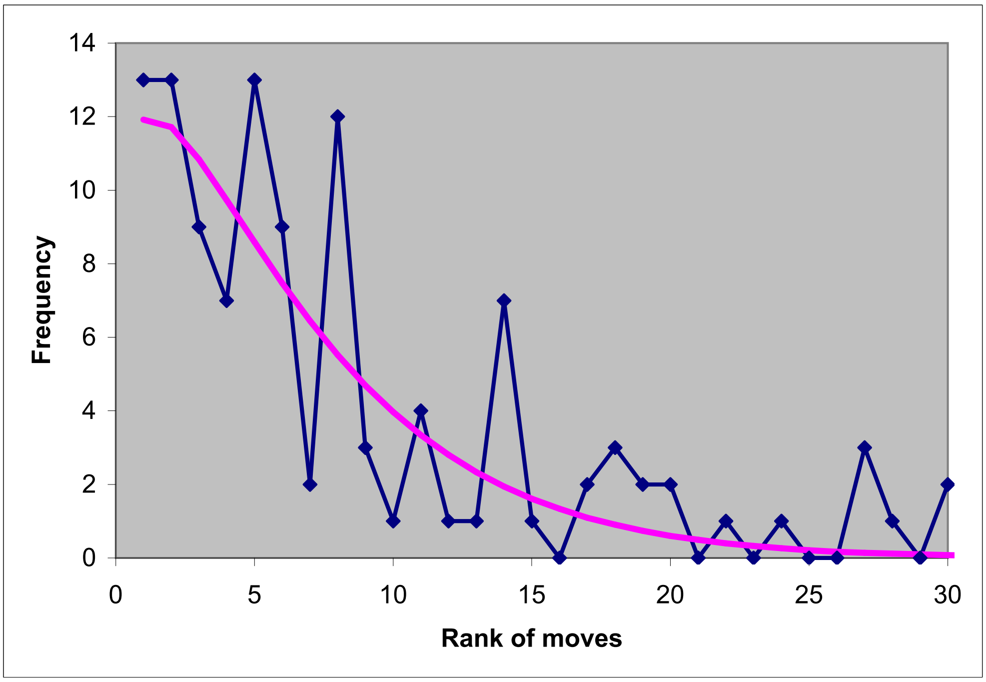

Figure 3-8: Frequency of moves at each rank, of 105 choreographed moves (line with markers) and Weibull approximation for frequency (solid line)

3.4.3 Results: Weibull fit to rank distributions

The raw distribution of move ranks within the 105 choreographed moves from two dance performance transcriptions appeared noisy. Rather than apply this distribution directly when invoking the automated choreographer, we fit a Weibull distribution to the entire dataset. The data are consistent with a Weibull (0.105, 1.17) distribution of ranks, as shown in Figure 3.4.3. This approximation, rounded to the next larger integer or rounded down to the maximum number of moves available from a given pose, is the actual distribution of move ranks that the automated choreographer employed.

3.4.4 Results: Musicality of choreographed versus extemporaneous dances

The minimization objective (3.2) can be used directly, to evaluate actual dance performances for their musicality. In twelve minutes of improvised dancing transcribed from the 2001 ALHC Jack and Jill contest, there were 235 moves with an average cost of 14.56 points per dance move. In the two transcribed dance routines which were choreographed specifically for the music, there were a total of 105 moves with an average cost of 14.82 points per move.

It is curious that these costs are nearly the same, when one might have expected to see a higher cost for the improvised dancing than for the carefully crafted routines. This particular Jack and Jill contest had an unusual feature: there were only two different music selections for all 11 couples, so most of the competitors got to hear their music in advance. All the dancers were advanced competitors with extensive experience at improvised dancing. As was described in Section 2.2, dancers can easily predict swing music's texture changes and moments of emphasis. It may be that for expert dancers, improvised dancing is in fact no less musically appropriate than pre-choreographed dancing. We will revisit the issue of musicality in improvised and pre-choreographed dancing, and consider the effect of possible communication errors in improvised dancing, in the next chapter.

3.5 Discussion

Creating an automated choreographer required modeling mathematically something that most viewers and indeed most dancers can conceive of only aesthetically. Because the dances considered here are carefully composed routines, each move likely has been considered carefully for its musical appropriateness.

The data fitting described here can only indirectly validate the optimization model that ranks alternatives. However, having attempted, and failed, to fit a deterministic optimization model exactly to the dances observed, we take the consistent bias (in the Kolmogorov-Smirnov test sense) towards highly-ranked moves as validation for the optimization model. Key inputs to the model were the numerical scores of dance moves and bars of music on attributes such as *exciting*. These inputs were constructed by the author based on extensive domain knowledge. Ranking alternatives lessens the dependence on the exact numerical values of the objective function, which seems proper when applying mathematical models to aesthetic judgments.

The primary challenge in fitting a model to data in this context is the variability in actual dance performances. Ultimately, the best an automated choreographer can hope for is the best a real dance couple could hope for: creativity in combining moves with the music that occasionally reaches the pinnacle of expression. It is said that the best swing dancers can make an audience hear things in the music that weren't noticed before. Merce Cunningham's quote bears repeating in this context; "Chance was a way of working which opened up possibilities in dance that I might otherwise have thought impossible" [19]. Our automated choreographer randomizes its choice of ranked alternatives in the hope of sometimes hitting the exactly right emotional chord. Its strategy is a balancing act between honoring the music and honoring the intelligence of the adversary, the audience and judges. Predictable dancing, like predictable war-making, can rarely achieve the highest success.

3.5.1 Applications

The obvious application for an automated choreographer is to create musically expressive animations of dancing characters that could replace the abstract visualizations provided in many computer music players. A number of expert Lindy Hop dancers contributed to this effort by dancing in a VICON motion capture studio to create source data for an animation. The VICON setup captured three-dimensional position data at 120 Hz for 41 markers each on two partners. The captured moves were scored on the attribute scales described in Section 3.3.1. Also, a few songs, distinct from the songs the dancers heard in the motion capture studio, were scored on the same scales. Visualization code created by Lukasz Hall rearranged the motion capture data to create animations of the automated choreographer's novel sequences of dance moves. An example of the final animation, showcasing the musicality of dances created by the automated choreographer, is in a video file accompanying this document [autodance.avi].

When surprise is valued, plans that optimize a particular objective may be inappropriate because the optimal actions could be predicted via inverse optimization methods [4]. The automated choreographer is a novel generator of close-to-optimal, yet unpredictable actions. A ranked alternative randomizer could be used to play games or to solve military resource allocation problems, the latter of which would make this model reminiscent of the origin of mathematical programming.

Further, the ranking method for determining how musically appropriate certain moves are could be used to objectively judge dance performances. Such a judge could not detect the musicality of the manner in which a move is performed, but only whether the moves chosen were broadly similar to the music. Indeed, if a traditional move is done in a very distinctive way then it might be appropriate to categorize the new version as different move. While a single fixed dance can not be the correct output of an optimal choreographer, there is certainly a difference between choreography which does and does not reflect the music, as demonstrated in Section 3.4 and in Chapter 2. An objective metric for this phenomenon could help train dancers and judges alike.

Chapter 4

The game of choreography

The previous chapter suggested that incorporating random move selections may be desirable for an automated choreographer, or for a human one. This chapter examines a game formulation of the choreographer's problem, in which the other game participant is a judge or an audience, to motivate that design choice. Randomizing over options with differing musicality scores is an optimal strategy for the dancers in the two-player zero-sum game described in this chapter.

We do not argue that dancers actually calculate a matrix for, and solve, a two-player game for every bar of music to be choreographed. Kahneman, in addressing whether humans are consistently rational, distinguishes between intuition, which is fast, effortless, and associational; and reasoning, which is slow, effortful, and rule-governed [67]. While choreographed routines certainly have been subjected to slow and effortful analysis, improvisational dancing is clearly outside the scope of direct reasoning behavior. Further, human decisions and judgements, even where deliberation is encouraged, deviate from purely rational behavior for a wide variety of tasks [96]. The development here supports the choice of randomization as a component of the automated choreographer. Validation is by analysis of the input - output similarity between the game model and the actual dance moves selected by experts for specific songs.

4.1 Predictability in music and dance

Music theorists have long held that tonal music creates schematic expectations in the listener, and that pleasing music must contain unexpected or surprising elements that sometimes violate those expectations [99, 66]. Likewise, dance performance should usually harmonize with the music but should also sometimes surprise the observer.

The information theory concept of entropy, a quantity that measures the unpredictability of a sequence of symbols, is a fitting way to measure the unpredictability of a dance. The dance language we considered for the entropy calculations consisted of 42 symbols.

First, we calculate the theoretical maximum entropy that could be realized by a dance language with the constraints on transitions that apply to Lindy Hop. The allowable transitions are shown in Figure 3-3. The entropy H of a first-order Markov process [3], with transition probabilities p_{ij} to state j from state i and steady-state distribution p_i, is:

$$H(s) = -\sum_i p_i \sum_j p_{ij} \log_2 p_{ij} \tag{4.1}$$

Assume the maximum entropy Markov probabilities, which make all allowable transitions out of every state equally likely and therefore achieve a uniform steady-state distribution of $p_i = \frac{1}{42} = 0.0238$. Let EC be the set of indices i of dance moves (states) that end in closed position, and let EO be the set of dance moves that end in open position. Further, let BC be the set of dance moves that begin in closed position, and let BO be the set of dance moves that begin in open position. Note that these four sets are not mutually exclusive; every move is in either BO or BC and also in either EO or EC. The only transitions which have non-zero probability are those between sets EC and BC and sets EO and BO. The cardinalities of these sets are 12, 30, 13, and 29, respectively. Then the theoretical maximum entropy of this subset of Lindy Hop is 4.49 by the following calculation.

$$H(s) = -\sum_{EC} p_i \sum s_{BC}(p_{ij} \log_2 p_{ij}) + \sum_{EO} p_i \sum s_{BO}(p_{ij} \log_2 p_{ij}) \tag{4.2}$$

$$= 12(\frac{1}{42})(13 * (\frac{1}{13}) \log_2(\frac{1}{13})) + 30(\frac{1}{42})(29 * (\frac{1}{29}) \log_2(\frac{1}{29})) = 4.49 \tag{4.3}$$

There are 1026 transitions in this matrix, and we have a corpus of only a few hundred moves. Rather than computing transition probabilities from an insufficient corpus, we calculate the entropy of the actual data from the observed single-move probabilities. The entropy, or unpredictability, of the choreographed dances here transcribed was 2.84, while the entropy of a transcribed collection of moves from improvised dances between people who were not regular partners was 1.86. Both transcriptions are from performances in the

American Lindy Hop Championships.

The choreographed and rehearsed dances were significantly less predictable than the improvised dances. It might be more difficult for the leaders to mentally access infrequently used moves during the extemporaneous performances. More likely, the leaders in an improvised setting may fear errors in communicating the more complex or less frequently used moves to an unknown partner. The leaders may choose to use a few basic moves more frequently, resulting in the lower entropy score for improvised dance. This latter argument will be formalized in Section 4.3 in a game matrix that considers the possiblity of communication errors.

4.2 Game theoretic model of dancers and judges

As an illustration, we draw a simple two-player zero-sum matrix game for a dance team and a judge. The dance couple may do a *swingout* to score 85 points or a *charleston* for 76 points. The judge guesses which move the couple might dance, and if the judge is correct he deducts 50 points for the predictability of their dance. The resulting payoff matrix for the dancers looks like

$$\begin{pmatrix} 35 & 85 \\ 76 & 26 \end{pmatrix}$$

Since this game contains no pure strategy saddle point, at the Nash equilibrium the dancers should use a mixed strategy and randomize its choice of move [28]. In fact, the dancers should do either a *swingout* or a *charleston* with equal probability of 0.5.

The score for performing a move is assigned based on the musicality distance metric of the previous chapter. Both Chapters 2 and 3 demonstrated that swing dancers place a high value on musicality. The number of points a couple will score with a particular move depends upon how well the move complements that bar of music. Then, one can view choreographing a song as a sequence of games, each with a different payoff matrix because each bar of music is different. Actually, choreographing a song is a multi-stage game because future payoffs are linked to present strategies, and this complete problem is outlined in Section 4.4 The score earned by move m in bar s in this section is $100 - d(m, s)$, where $d(m, s)$ is as in (3.1).

Other factors that might contribute to a judge assigning a move score are the move's difficulty and the intangible style or grace with which the move is executed. The dancers' grace can be assumed to apply to each move equally, so that component of the score has little influence the choice of choreography. The difficulty of a move might be important to scoring, but we neglect that component of judges' scores.

Consider the leader's choice of a move for bar 12 of the song Frenesi. The move danced in bar 11 ends in closed position. There are 13 moves that start in closed position to choose from for bar 12. The moves and their musicality scores are shown in the first three columns of table 4.1. The judge also knows the vocabulary of moves and the structure of the music, and the judge again deducts 50 points if he guesses correctly which move the leader will dance. The one-step zero-sum payoff matrix for dancing to this bar of music is

$$\begin{pmatrix}
-12 & 38 & 38 & 38 & 38 & 38 & 38 & 38 & 38 & 38 & 38 & 38 & 38 \\
91 & 41 & 91 & 91 & 91 & 91 & 91 & 91 & 91 & 91 & 91 & 91 & 91 \\
76 & 76 & 26 & 76 & 76 & 76 & 76 & 76 & 76 & 76 & 76 & 76 & 76 \\
25 & 25 & 25 & -25 & 25 & 25 & 25 & 25 & 25 & 25 & 25 & 25 & 25 \\
78 & 78 & 78 & 78 & 28 & 78 & 78 & 78 & 78 & 78 & 78 & 78 & 78 \\
85 & 85 & 85 & 85 & 85 & 35 & 85 & 85 & 85 & 85 & 85 & 85 & 85 \\
74 & 74 & 74 & 74 & 74 & 74 & 24 & 74 & 74 & 74 & 74 & 74 & 74 \\
71 & 71 & 71 & 71 & 71 & 71 & 71 & 21 & 71 & 71 & 71 & 71 & 71 \\
69 & 69 & 69 & 69 & 69 & 69 & 69 & 69 & 19 & 69 & 69 & 69 & 69 \\
85 & 85 & 85 & 85 & 85 & 85 & 85 & 85 & 85 & 35 & 85 & 85 & 85 \\
90 & 90 & 90 & 90 & 90 & 90 & 90 & 90 & 90 & 90 & 40 & 90 & 90 \\
70 & 70 & 70 & 70 & 70 & 70 & 70 & 70 & 70 & 70 & 70 & 20 & 70 \\
88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 38
\end{pmatrix}$$

and the optimal strategy for the leader is to choose one of the moves: *closedfootwork, charleston, flashkicksjockey, sendout, swingout*, or *tuckturn*; each with probability $\frac{1}{6}$. This is a maximum entropy distribution over the moves with musicality scores greater than or equal to 76. The optimal strategy for any of these choreography games will always be of a simple form: Choose each of the moves that is "good enough" with equal probability. The moves which are good enough are the g highest-scoring moves, where g is defined implicitly by inequality to be introduced later 4.9. Just as infinite video loops try in a sense to maximize information flow to a viewer [32, 122], so the dancers try to maximize the information communicated by their dance. The following theorem gives the general solution for the above-described choreography game.

Theorem: Two-player zero-sum games with a payoff matrix of the form

$$\begin{pmatrix} a_1 - j & a_1 & \dots & a_1 \\ a_2 & a_2 - j & \dots & a_2 \\ \vdots & \ddots & \ddots & \vdots \\ a_n & a_n & \dots & a_n - j \end{pmatrix} \tag{4.4}$$

have optimal row strategy that uses the g moves with highest a_i, each with probability $\frac{1}{g}$, where g satisfies inequality (4.9).

Proof: Without loss of generality, order the rows and columns by increasing score a_i. The proof is by the matrix form of the simplex algorithm on the column player's problem. The penalty to the dancer if the judge guesses correctly is j, and let $y_1, \dots, y_n$ be the judge's (column) strategy. Add $y_{n+1}, \dots, y_{2n}$ slack variables to get the program: Minimize M, subject to

$$\begin{array}{rrcrrcc} (a_1 - j)y_1 & +a_1 y_2 & \dots & +a_1 y_n & +y_{n+1} & = & M \\ a_2 y_1 & +(a_2 - j)y_2 & \dots & +a_2 y_n & +y_{n+2} & = & M \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_n y_1 & +a_n y_2 & \dots & +(a_n - j)y_n & +y_{2n} & = & M \\ y_1 & +y_2 & \dots & +y_n & & = & 1 \\ y_1, & y_2, & \dots, & & y_{2n} & \geq & 0 \end{array} \tag{4.5}$$

where M is the maximum dancer (row) payoff. We first re-order the variables slightly and perform a few elimination steps to reach the tableau form from [28], page 292:

$$
\begin{array}{r|cccccccccc}
\text{Constants} & y_n & y_{n+1} & \dots & y_{2n-1} & y_{2n} & y_1 & y_2 & \dots & y_{n-1} \\
\hline
1 & 1 & 0 & \dots & 0 & 0 & 1 & 1 & \dots & 1 \\
a_n - a_1 - j & 0 & 1 & \dots & 0 & -1 & -2j & -j & \dots & -j \\
a_n - a_2 - j & 0 & 0 & \ddots & 0 & -1 & -j & -2j & \dots & -j \\
\vdots & & & \ddots & & \vdots & & & \ddots & \\
a_n - a_{n-1} - j & 0 & 0 & \dots & 1 & -1 & -j & -j & \dots & -2j \\
\hline
\text{Cost } c & 0 & 0 & \dots & 1 & 0 & j & j & \dots & j
\end{array}
\tag{4.6}
$$

The first column of the tableau is referred to as b, the row below the lower line as c, and the remainder of the tableau above the lower line as P. Consider the basis B consisting of the columns $\{y_n, y_{n+1}, \dots, y_{n+r}, y_{r+1}, y_{r+2}, \dots, y_{n-1}\}$ of P, with r being the number of moves the dancers will *not* dance. Then $g = n - r$ is the number of moves the dancers will choose from, and g takes integer values between 1 and n. Let γ be the components of c corresponding to the basis columns. Recall that the basis B is feasible if $B^{-1}b \geq 0$, and optimal if the reduced costs $\bar{c} = c - \gamma B^{-1} P \geq 0$.

Then B^{-1} has the block structure

$$
B^{-1} = \left[\begin{array}{cccc|cccc}
1 & 0 & \dots & 0 & \frac{1}{gj} & \frac{1}{gj} & \dots & \frac{1}{gj} \\
0 & 1 & \dots & 0 & -\frac{1}{g} & -\frac{1}{g} & \dots & -\frac{1}{g} \\
\vdots & & \ddots & \vdots & \vdots & \vdots & \vdots & \\
0 & 0 & \dots & 1 & -\frac{1}{g} & -\frac{1}{g} & \dots & -\frac{1}{g} \\
\hline
0 & 0 & \dots & 0 & -\frac{(g-1)}{gj} & \frac{1}{gj} & \dots & \frac{1}{gj} \\
0 & 0 & \dots & 0 & \frac{1}{gj} & -\frac{(g-1)}{gj} & \dots & \frac{1}{gj} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \dots & 0 & \frac{1}{gj} & \frac{1}{gj} & \dots & -\frac{(g-1)}{gj}
\end{array}\right]
\tag{4.7}
$$

where the dimension of the upper left block is (r+1)-square and the dimension of the lower right block is (g-1)-square. This structure enables closed-form calculation of the required quantities. The reduced costs are positive, because

$$
\bar{c} = [\overbrace{0 \quad \dots \quad 0}^{r+1} \quad \overbrace{\frac{1}{g} \quad \dots \quad \frac{1}{g}}^{g} \quad \overbrace{\frac{j}{g} \quad \dots \quad \frac{j}{g}}^{r} \quad \overbrace{0 \quad \dots \quad 0}^{g-1}]
\tag{4.8}
$$

so the basis B proposed is optimal, provided that it is feasible. The $2, 3, \ldots, (n+1)$ components of $\bar{c}$, which are those corresponding to variables $y_{n+1}, \ldots, y_{2n}$, are the optimal probabilities for the n moves in the dancers' (row player's) strategy [28]. Each of the g highest-scoring moves is equally probable.

The basis B is feasible if g can be set to satisfy $B^{-1}b \geq 0$. Note a_{n-g+1} is the score of the lowest-scoring move that will be chosen with non-zero probability in the optimum row strategy. Call the differences between the best move and each of the other moves $\Delta_k = a_n - a_{n-k}$, so that Δ_k increases with k, $0 \leq \Delta_1 \leq \Delta_2 \leq \ldots \leq \Delta_{n-1}$. Interpreting the $\geq$ sign component-wise, the simplified set of inequalities to be satisfied is

$$\begin{bmatrix} \frac{1}{g} + \frac{1}{gj}[\Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \\ \Delta_{n-1} - \frac{1}{g}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \\ \Delta_{n-2} - \frac{1}{g}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \\ \vdots \\ \Delta_g - \frac{1}{g}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \\ -\frac{1}{j}\Delta_{g-1} + \frac{1}{gj}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \\ -\frac{1}{j}\Delta_{g-2} + \frac{1}{gj}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \\ \vdots \\ -\frac{1}{j}\Delta_1 + \frac{1}{gj}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \end{bmatrix} \geq 0 \tag{4.9}$$

All terms in the first component of (4.9) are positive, so the first component is positive. The parallel forms of the following r terms of (4.9) guarantee that all of these terms are positive if the last one is positive, $\Delta_g - \frac{1}{g}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \geq 0$. Likewise the last $g-1$ terms in (4.9) will be positive as long as the first one is positive,$-\frac{1}{j}\Delta_{g-1} + \frac{1}{gj}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}] \geq 0$.

Let $f(g) = \frac{1}{g}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{g-1}]$, so $f(g)$ is the average of j and the $g-1$ differences betwen the best move score and the $g-1$ next best dance move scores. Then the inequalities to be satisfied can be simplified to

$$\Delta_{g-1} \leq f(g) \tag{4.10}$$

and

$$\Delta_g \geq f(g) \tag{4.11}$$

At the extreme values of g, if $g = 1$ only $\Delta_1 - j \geq 0$ (4.11) must be satisfied, and if $g = n$ then only $\Delta_{n-1} \leq \frac{1}{n}[j + \Delta_1 + \Delta_2 + \ldots + \Delta_{n-1}]$ (4.10) must be satisfied.

There will always exist at least one g that satisfies the required inequalities. If only one move is to be included in the strategy, $g = 1$, then $\Delta_1 \geq j$, and so the second-best move's score must be more than j points away from the best move's score. By definition of an average, if a number Δ is smaller than the average of a set K of numbers, then the average of set $K \cup \{Delta\}$ is smaller than the average of set K. We will use this property in what follows. By finite induction on $k \in \{1, 2, \ldots, n-1\}$, there will be a valid g. It is easy to show that if (4.11) is false for $g = k$, then (4.10) will be true for $g = k + 1$. If (4.11) is false for $g = k$ then $\Delta_k < f(k)$. Then, (4.10) must be satisfied for $g = k + 1$ because the average $f(k + 1)$ now incorporates Δ_k and Δ_k was smaller than the previous average $f(k)$. At the last step, if $k = n - 1$ does not yield a valid g, then the inequality (4.10) for $g = n$ will have been satisfied, and since the inequality (4.11) for $g = n$ does not apply, $g = n$ satisfies (4.9).

We do not argue uniqueness of g; there may be degenerate cases in which more than one g satisfies the inequalities, in which case there are multiple solutions to the LP and multiple optimum row strategies.

4.3 Communication errors in improvised dancing

The leader's choice of choreography in an improvisational Jack and Jill contest is also likely to be influenced by the relative difficulty of communicating various moves. In a Jack and Jill contest, not only are the dance moves chosen on the fly, but the partners might never have danced together before. Table 4.1 shows the probabilities of an error for the moves feasible at the 12th bar of music. The probabilities and penalties in Table 4.1 are based on the author's Jack and Jill and social improvised dance experience.

When errors of this sort do happen, the penalty is relatively large. The penalty also depends upon the particular move, because some moves are unlikely to go drastically wrong while others will look very awkward if they don't go smoothly. The product of the probability of and the penalty for an error is the expected number of points the couple will lose for miscommunications that are visible to the judges.

After accounting for the expected penalty for miscommunications, the game matrix for improvised dancing to the 12th bar of the song Frenesi is

Table 4.1: Move scores and probability, cost, and expected penalty for communication errors

Move	Score	p(error)	c(error)	E(penalty)
foxtrot	38	0.01	800	8
closedfootwork	91	0.04	500	20
popturntoclosed	76	0.04	500	20
Balboa	25	0.02	200	4
charleston	78	0.005	200	1
flashkicksjockey	85	0.01	500	5
freeze	74	0.02	800	16
eastcoast	71	0.005	400	2
promenade	69	0.01	500	5
swingout	85	0.01	400	4
sendout	90	0.01	500	5
popturn	70	0.03	500	15
tuckturn	88	0	100	0

$$\begin{pmatrix}
-10 & 30 & 30 & 30 & 30 & 30 & 30 & 30 & 30 & 30 & 30 & 30 & 30 \\
71 & 31 & 71 & 71 & 71 & 71 & 71 & 71 & 71 & 71 & 71 & 71 & 71 \\
56 & 56 & 16 & 56 & 56 & 56 & 56 & 56 & 56 & 56 & 56 & 56 & 56 \\
21 & 21 & 21 & -19 & 21 & 21 & 21 & 21 & 21 & 21 & 21 & 21 & 21 \\
77 & 77 & 77 & 77 & 37 & 77 & 77 & 77 & 77 & 77 & 77 & 77 & 77 \\
80 & 80 & 80 & 80 & 80 & 40 & 80 & 80 & 80 & 80 & 80 & 80 & 80 \\
58 & 58 & 58 & 58 & 58 & 58 & 18 & 58 & 58 & 58 & 58 & 58 & 58 \\
69 & 69 & 69 & 69 & 69 & 69 & 69 & 29 & 69 & 69 & 69 & 69 & 69 \\
64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 24 & 64 & 64 & 64 & 64 \\
81 & 81 & 81 & 81 & 81 & 81 & 81 & 81 & 81 & 41 & 81 & 81 & 81 \\
85 & 85 & 85 & 85 & 85 & 85 & 85 & 85 & 85 & 85 & 45 & 85 & 85 \\
55 & 55 & 55 & 55 & 55 & 55 & 55 & 55 & 55 & 55 & 55 & 15 & 55 \\
88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 88 & 48
\end{pmatrix}$$

The solution to the above game is that the leader should choose one of the moves: *charleston, flashkicksjockey, sendout, swingout*, or *tuckturn*; each with probability $\frac{1}{5}$. A similar strategy was optimal for the pre-arranged dance game in Section 4.2, except that for the pre-arranged dance the move *closedfootwork* was also included in the leader's set. The error penalty and probability of error for the move *closedfootwork* are high, so the leader should not consider this move for an improvised dance even though he would consider it for a choreographed dance.

It is not clear whether this communication penalty will result in a dance with lower entropy, as would be expected from the difference discovered between the entropies of actual choreographed and improvised dances in Section 4.1. To define a condition on the musicality scores and communication penalties which would guarantee this inequality holds is left to future work.

4.4 Multi-stage choreography game

Choreographing an entire song with M bars is a finite multi-stage game with one stage per bar of music. Due to the move continuity restrictions, future payoffs depend upon the decisions made at every stage. It is not sufficient to solve the one-step game as described in Section 4.2 at each bar of music. The multi-stage game can be solved with a dynamic programming approach by working backward from the final stage at the final bar of music [7, 9]. Let the one-step payoff matrix for each bar i of music where the previous move was k be $R_{i,k}$. Let the expected future payoff from stage i to stage n if the move at stage i is k be $V_{(i)}(k)$. Then $V_i(k)$ is the value of the game with matrix

$$R_{i,k} + \begin{pmatrix} V_{i+1}(1) & V_{i+1}(1) & \dots & V_{i+1}(1) \\ V_{i+1}(2) & V_{i+1}(2) & \dots & V_{i+1}(2) \\ \vdots & \vdots & & \vdots \\ V_{i+1}(n) & V_{i+1}(n) & \dots & V_{i+1}(n) \end{pmatrix} \tag{4.12}$$

Note that because $R_{i,k}$ has the form (4.4), the payoff matrix (4.12) also has exactly the same form (4.4). Calculating the optimal strategy for the game with payoff matrix (4.12) is a simple linear program provided one can calculate $V_{i+1}(k), \forall k$. This multi-stage game is easily solved by backward recursion, starting with the initialization step that $V_{(M+1)}(k) = 0, \forall k$.

4.5 Game model versus optimization model

Both the game model of this chapter and the optimization model discussed in Chapter 3 assigned scores to dance moves for their musical appropriateness, and both explicitly require

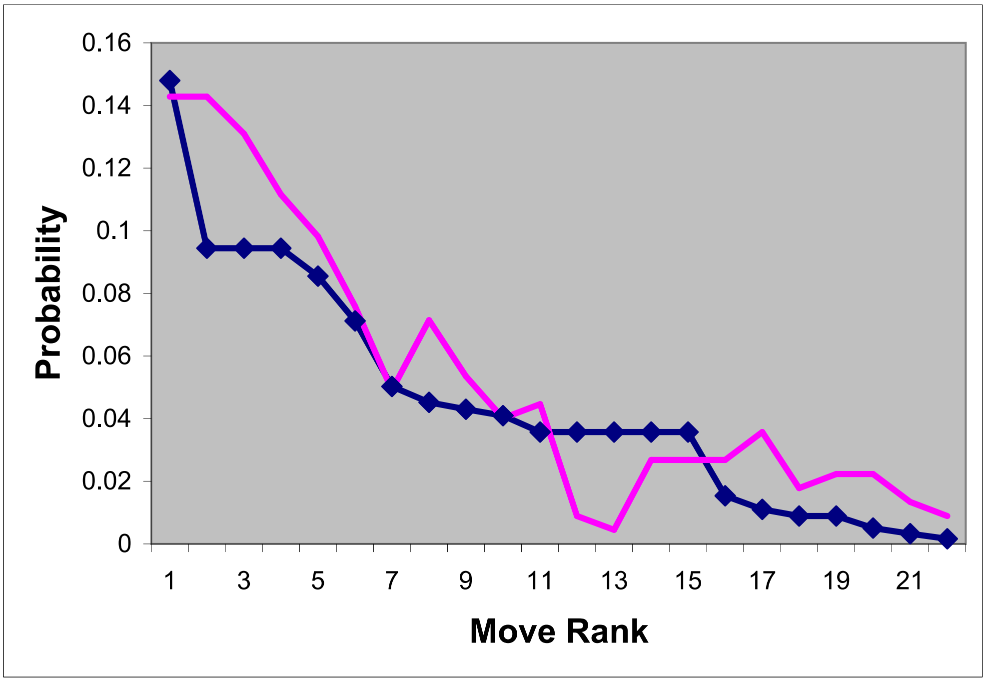

Figure 4-1: Smoothed probability of actual ranks of moves (solid line) compared with game theory model prediction if good-enough moves are equally probable (line with markers), for the song *Frenesi*.

some random element in the choice of moves. At first glance, though, the game model's predicted uniform distribution over moves that are "good enough" seems incompatible with the biased distribution over ranked dance moves from Chapter 3.

However, recall that a different game matrix will be generated for each bar of music. Perhaps some bars will have the dancer choosing with equal probability from the 20 best moves while other bars of music will have him choosing with equal probability from only the 3 best moves. Then the overall probability of choosing moves with particular ranks would appear biased towards highly-ranked moves. As an illustration of this, we calculated the number of good-enough moves, moves within 20 points of the highest ranked move, for every bar of music and starting pose for the actual *Frenesi* dance. Assigning a uniform distribution over the number of good-enough moves and summing over the 56 bars of music in the song, we found a distribution biased toward highly-ranked moves that appears similar to the actual distribution over the move ranks, as shown in Figure 4.5.

4.6 Conclusion

The game model of musically expressive and entertaining choreography presented here is admittedly simple. There are many aspects of dance and judging which have been neglected. Further, showing that a model can predict reality, even if it does so perfectly, does not prove that the model describes the actual mechanism that is operating. Still, this investigation of dancing for judges offers a model in which leaders, making rational decisions, will randomize their choice of move rather than select the single optimal dance move.

The dancer takes an adversarial relationship to the judges or the audience, in that the dancer always aims to surprise while remaining musically expressive. Military planning problems, the subject of decades of sustained research, likewise have a primary objective with the side objective of achieving surprise. The wide applicability of a good decision-maker that evades prediction makes modeling choreography a difficult but worthwhile enterprise.

Chapter 5

Robot leader experiments

Partner dancers can coordinate actions between human leaders and followers, who share a rhythmic motion vocabulary, by haptic (touch) signaling with an audio metronome. This chapter asks whether human-robot cooperative tasks might benefit from the use of rhythmic motion vocabularies similar to those found in human-human dance interaction. To determine the effect of a shared motion vocabulary in a dance-like task, this study uses a PHANToM haptic robot to lead a dance followed by human subjects. The data demonstrate that the subjects' understanding of the motion as a random sequence of known vocabulary elements allows them to follow using prediction of the lead. This type of interaction is fundamentally different from closed-loop pursuit tracking because the subject could at least in principle recognize the current state of the high-level automaton, as described in Chapter 3. We found that subjects do use their knowledge to react to upcoming input, defeating the physiological minimum delay for closed-loop pursuit.

5.1 Background, and relevance of human-robot dance

The leader and follower dancing model is a useful instance of the multi-agent systems framework for cooperative tasks between man and machine [81]. Observations from the successful haptic collaboration system that exists between dancers can be used in designing human-robot haptic collaboration. Other researchers are interested in human-robot collaboration in a dance setting: Setiawan et al. have designed a robot that follows a walking human [123], and Kazuhiro Kosuge's group at Tohoku University created a ballroom dance follower robot on a wheeled platform [75].

Human-robot haptic cooperation tasks are tasks that place the human and robot in contact with each other, or enable both to exert forces on a shared tool. This class includes strength-extending wearable robots [71] and teleoperator systems [125, 126].

Teleoperation systems have a master device operated by a human and a remote slave device which manipulates an environment [125]. They are useful in domains where the human would not directly perform the task because it is either too far away, too dangerous, or at too small a scale, or has too restricted a workspace as in minimally invasive surgery [107]. Creating telepresence, the sense that the user is actually performing the task, has usually taken the form of reflecting forces between the slave device and its environment to the master device [83]. Bilateral teleoperation refers to the various forms of feedback from the environment (position feedback, force feedback, impedance reflection) that use the master to impose forces on the user. Haptic feedback has been shown to reduce tissue-damaging errors in a telesurgical dissection [138]. Haptic virtual reality systems have a similar goal of presenting forces to the user, although in this latter case the environment is only simulated [56]. Only the kinesthetic (force) component of haptic information is communicated in the vast majority of systems, as programmable tactile (texture) displays are not yet practical.

Bilateral teleoperation and haptic virtual reality are focused on providing kinesthetic feedback to the user, which is exactly the type of information which we will argue is valuable to swing dancers in the structured task of dancing. A virtual fixture is a simulated addition to a teleoperated environment that can aid performance of a skilled task [91, 120]. The experiment reported here demonstrates the usefulness of a specific style of touch cues, which suggests a novel virtual fixture that leads the proper motion primitive from a known set. A dancing game with the robot as leader could also add excitement and interactivity to repetitive movement robot rehabilitation [114].

Bilateral teleoperation is notoriously problematic because of the possibility of energy being injected by the master and slave controllers [72], with the seemingly obvious position-forward force-feedback approach having serious stability problems when the slave contacts a rigid environment. Most resolutions of the stability problem have utilized the notion of passivity to prevent the (perhaps virtual) environment and devices from adding energy to the master [56, 2, 83]. Passivity is not a reasonable assumption for swing dance partners, who depend upon a transfer of energy to complete their movements. In fact, the PD control loop which ran the PHANToM haptic device as a dance leader in the present experiment is active, injecting energy to move the participants' hands. The energy added was not great enough to enable participants to perform acceptably by letting the PHANToM drag their hands around.

Human-human contacts do not exhibit the stability problem seen in teleoperation systems, because human limbs are in fact tuned to act as passive impedances [59, 102]. Although there is no actual delay in the mechanical interaction between two human limbs, there is a long-loop delay for each person before cognitive processing can take place and muscle tension settings can be adjusted. Wave variables, which enable graceful degradation of performance with delay in teleoperation [104], have been theorized to perform the same function in the human motor control system [92]. One of the conclusions of the present experiment is that in dance-like tasks that allow communication of the upcoming vocabulary input, a person's response to kinesthetically-communicated direction reversal need not wait for the long-loop delay.

Our PHANToM robot plays the role of the leader with a human subject as follower. A robot leader with a human follower could act as a teacher [37, 49], or as a virtual experience guide for playback of haptic or touch recordings, as in [142]. The roles of leader and follower are not always sharply drawn, because dance leaders will often adjust to accommodate a follower's style or a mistake made. In dance, followers occasionally take charge to decide what move will come next [46].

In fact, rhythm, timing, and to a lesser extent move vocabularies have often been neglected in human-robot interaction work. Dancers embody a gentle interaction, in which each participant extends the capabilities of the other by cooperation facilitated by a shared understanding of the motions to be performed. Such concepts could make robot assistants more welcome, for instance, in the operating room or the living room.

Human-robot team surgery is a meaningful prospective application for rhythmic motion vocabularies. Surgeons use telemanipulation to improve surgical outcomes and minimize incisions. Further, surgery requires precision movements that have been categorized, for example, into grasping, spreading, and sweeping tissue [119], and the choices made among these movements correlate to surgical skill. A robot assistant might train novice surgeons to choose the proper sequence of actions in a simulated surgical environment.

The daVinci surgical robot, in use today, enables surgeons to teleoperate tiny laparoscopic instruments with articulated wrists. Currently the device uses no force feedback [54]; miniaturized, tissue-compatible, and disposable 6-axis force sensors are prohibitively expensive [107]. Fictional force feedback from a purely visual force sensing system has recently been proposed [73]. Literal force sensing for feedback might be replaced by force feedback from virtual fixtures [120], if registration can be resolved and fixtures which improve performance of real surgical tasks can be designed. With the surgeon as leader, a virtual fixture follower might follow expert surgeons through a procedure to compensate for a lack of haptic feedback and speed up sewing, prevent injury of vital nerves or vessels, or reduce hand jitter [128], all without removing control from the surgeon and conditioned on the recognition of the surgeon's current movement. An intent recognition system based on hidden Markov models is described in Li and Okamura [86].

5.2 Overview

Partnered dancing requires coordination of actions between two humans by haptic, or touch-based, communication. Previous work on haptic interfaces between a cooperating pair of humans found that haptic-only cooperation was inferior to visual-only cooperation and found that haptic-plus-visual cooperation was not superior to visual-only cooperation. The task was for subjects to teach one another how to draw an unfamiliar Japanese calligraphic character using a pair of bilateral teleoperated PHANToM haptic devices [106]. However, experienced swing dance followers have demonstrated in our lab the correctly identify moves and to follow them while blindfolded. The critical difference between these tasks is the highly structured nature of swing dance. Swing dancers share a synchronizing audio grid and a rhythmic move vocabulary, with one of these known moves assigned to each bar of music. Dance interactions appear more robust to visual deprivation than less structured cooperative tasks.

One goal of this study was to demonstrate improved human-robot interaction by restriction of the cooperative task to sequences of known moves. This study is primarily to investigate the participants' interaction strategies in the specific context of improvised dance, but also aimed to test general ideas about the nature of human-robot haptic communication. Our results show that delay before responding to robot cues is reduced with known moves. The study investigates whether dance followers use mental models of each move.

We hypothesize that dance followers will recognize each move as it is led, and that recognition will enable some open-loop or predictive control. The primary metrics used to determine whether followers have recognized a move are the length of time delays and the frequency of position errors. Time delays and position errors are measured at a direction reversal within the dance pattern. A direction reversal is where a clockwise tracing of a circle switches to counterclockwise tracing of that circle, or vice versa. Reversals are either predictable if the move has been recognized, or unpredictable if the reversal is at one of the random transitions between moves. If dance followers do not use their knowledge of the vocabulary of dance moves, then none of the reversals should be predictable and there should be no difference between the time delays and position error frequencies observed within moves and at move transitions.

Hypothesis 3 *Subjects in the dance following experiment will use open-loop or predictive strategies when they have recognized an element from the dance move vocabulary. Subjects' performance at direction reversals, measured by time delay and position error frequencies and variance in position traces, will therefore be different depending upon whether the direction reversal was predictable as part of a move or unpredictable at a move transition.*

5.3 Prior experiments

5.3.1 Manual tracking

In the literature on manual tracking tasks, central nervous system delays on the order of 160 ms have been observed [34]. Long-loop delays have a strict lower limit of about 100 ms, because nerve impulses require a finite travel time to and from the brain. The short-loop stretch reflex of muscle can have a shorter delay, on the order of 60 ms. The length of long-loop central nervous system delays make for poor performance of proportional control strategies. Fuzzy controllers, which have some commonality with the dance move recognition postulated here, can reduce settling time when delays are of the order of central nervous system delays [51].

A brief review of a number of classical results in control system modelling of human operators can be found in Hess [57]. These results are formulated for random or random-appearing inputs or disturbances. A random sequence of dance moves is not a random-appearing input, because the moves are known and transitions between moves are constrained to happen at specific times during a dance. Move transitions land on the beat of the music that both partners hear.

More closely related to the current study are Jagacinski's experiments [63]. His subjects repeatedly tracked identical 20 second segments with a position control joystick via visual display of error. With practice, subjects were able to reduce the effective delay of their responses from about 126 ms to about 32 ms. Effective delay was calculated as the best fit τ in a Taylor-series approximation to an input reconstruction model $x(t) = \beta_0 + I(t - \tau)$, where $x(t)$ is the subject's position and $I(t - \tau)$ is the input τ seconds prior. As 32 ms delays in visuomotor loops are physiologically infeasible, this effective delay reflects learning of the input signal I.

Then, the current experiments can be seen as exploring the space between exactly repeated tracking tasks and apparently random input tracking tasks. In this work, as in extemporaneous partner dance, repeated moves are sequenced in either novel or predictable ways.

5.3.2 Haptic modality

Classical research on human tracking ability considered mostly tracking visual displays of error using a joystick. The presentation of error here is novel. While the subjects can see the path of the PHANToM's stylus, they can only perceive error levels haptically, interpreting greater forces as greater errors. It is to be expected that position error is higher in this haptic-display movement task than it would be if the task included a visual error display, because the fidelity of the haptic sensing system is lower. Weber's law holds that just noticeable difference (JND) is a pure number, the percent change in the magnitude of any signal that will be reliably perceived as different from the standard signal. The traditional JND=0.02 for visual discrimination of luminance. Just noticeable differences are larger for kinesthetic signals than for visual ones, with JND = 0.10 [5]. Humans can maintain what they believe to be a constant force with their elbow joint only to within a mean error of 4.5N with haptic feedback, but can improve that to 0.76N with visual plus haptic feedback [65].

Although visual display is superior to haptic display, there will be applications in which it is best to minimize visual demands upon the user. Jagacinski et al. compared visual display of error to tactile display of error and found visual display to be generally superior [62]. Jagacinski et al. used a different type of haptic display than the one considered here. Their device rolled across the subject's fingertip to indicate position errors, or combined position and velocity errors, but the PHANToM device instead imposes forces that contribute to the motion of the stylus.

Wagner and Howe distinguished between force feedback as a physical constraint on motion and force feedback as a means of displaying information [137]. They compared reversals of direction when subjects encountered a virtual wall implemented as either: a spring of varying stiffness, or a vibration which transferred no energy. Subjects could not respond to the vibratory wall until after a 150 ms delay, but the spring force acted as a physical error correction that reduced the subjects' incursion into the virtual wall. Wagner and Howe also use a second-order model of the hand and arm to estimate the subjects' desired trajectory absent the physical interaction with the force. They conclude that higher spring constants not only reject the incursions physically, but also reduce the time delay to reversing the subjects' desired trajectory.

Our experiment instead compares apples to apples, that is, both the predictable and unpredictable direction reversals in our experiment use the same controller and reference trajectory. We investigate whether the subjects' performance in these two categories differs, to show whether subjects are actively predicting reversals within the randomly sequenced dance moves. Unlike Wagner and Howe [137], we do not separate the following possible subject strategies: (1) setting certain parameters of the limb differently for a predicted direction reversal so that the PHANTOM forces the limb in the reverse direction, and (2) incorporating the predicted direction reversal into the desired trajectory.

5.3.3 Successive organizations of perception and internal models

The successive organizations of perception theory of humans as controllers describes three levels of behavior: compensatory, pursuit, and precognitive [90]. In compensatory tracking, the subject's control input is the perceived error. In pursuit tracking, the subject's control inputs are perceived error and perceived or estimated target. Even if, as in this experiment, the target is not displayed, experienced subjects estimate it from the displayed error and behave as if it were present [63]. In precognitive tracking, the subject mentally describes a pattern for the target and tracks in a precognitive, or open-loop, fashion.

The question this study asks is whether the subjects will use precognitive controls when presented with repeated moves sequenced unpredictably.

Neuroscience researchers have investigated internal model hypotheses about how the motor control system predicts consequences of motor commands. In particular, Karniel and Mussa-Ivaldi trained people on two novel force environments, and subjects were able to gain control of their movements under these force fields with practice. However, their subjects were unable to maintain both models and switch between them on successive trials, even though the simple alternation of models was predictable [70]. Our study differs from Karniel and Mussa-Ivaldi's in that our subjects are exposed to a random sequence of known moves or trajectories, and our subjects probably do not require novel central nervous system models of motor control to follow these moves. We found evidence that our subjects were able to adjust their intended trajectories after recognizing the move.

5.4 Experiments

The PHANToM 1.5, a commercial haptic device shown in Figure 5-1, presents a force at its stylus endpoint which is a function of its sensed position. The PHANToM is a very high-end desktop device which is capable of 1000 Hz precision force-feedback [93]. The device provides the Cartesian coordinates of the end-effector through the software development kit. The programmer computes the force to be output at the end-effector. There are no force encoders at the end-effector. The force commanded is the only force that can be recorded, and is assumed to represent the interaction force although these are probably not identical since the device is in motion. Other haptic display devices are available [134]. This study used the PHANToM to implement a dance-like game with the robot as leader and human as follower [47, 48].

The PHANToM leads a human follower in unknown sequences of known moves. The device follows a trajectory with a modified proportional-derivative (PD) feedback controller. The force F imposed on the end-effector, is a function of the error signal e_p and e_v in x and y, the differences respectively of position and velocity of the stylus from the reference position and velocity. The moves are two-dimensional circular and semi-circular shapes in x and y, with z controlled but not constrained to a fixed value. Here, positive x is the direction to the right of the user, positive y is the direction towards the ceiling, and positive

Figure 5-1: PHANToM haptic device

z is perpendicular to x and y axes and pointing toward the user.

The control force which the user feels is

$$F = k_z e_z + (1 + k_g |e_z|)(k_p e_p + k_v e_v) \tag{5.1}$$

with $k_z = 0.03$N/mm, $k_g = 0.01$N/mm, $k_p = 0.005$N/mm, $k_v = 0.0005$N-s/mm. The first term enforces a preferred plane for the moves, and the second term is the PD feedback, with the overall gain increasing with distance from the preferred plane due to the k_g term. The justification for increasing the gains with distance from the preferred plane is addressed in Section 5.5.4.

Using these gains, the forces imposed are gentle and could not drag a subject's hand through the moves without the subject's active participation. Although the device's dynamics surely influenced the hand's dynamics as in Wagner and Howe's experiment [137], the haptic dance is designed to encourage active participation and, if possible, prediction. The forces imposed were everywhere less than .7N, and averaged about .24N. With no user touching the robot as it led these moves, the stylus will move in space, but the feedback gains are so small that the total motion covers a circle smaller than the reference trajectory (see Figure 5-2). If a user failed to add his own strength to the movements led but added mass to end-effector, these too-small circles would have become even smaller. None of the subjects used this passive strategy.

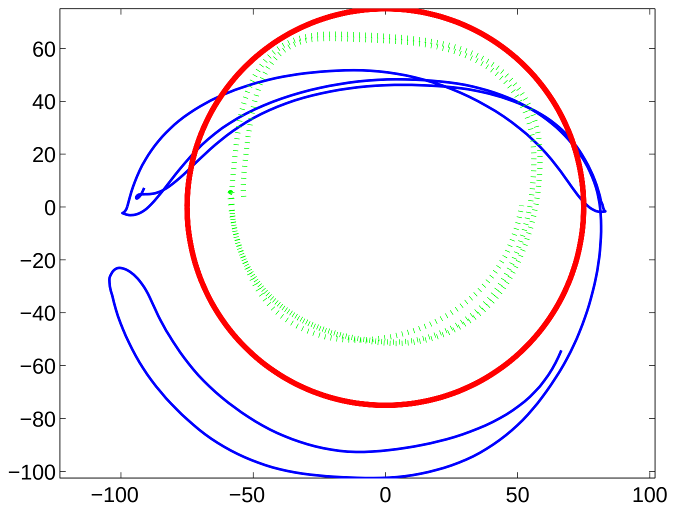

Figure 5-2: Without user intervention (dotted line), the circles traced were smaller than the reference trajectory (thick solid line). With user interaction, the circles traced were larger (thin solid line).

5.4.1 Dance vocabulary

The haptic dance has syntactic content, as do the standard forms of partner dance. Moves are selected from a set of four moves: two clockwise circles, two counterclockwise circles, four upper half circles, and four lower half circles, all of which are executed in the frontal plane. The frontal plane goes left-right and up-down, excluding toward-away from the subject. The moves are illustrated in Figure 5-3.

The reader can view an MPEG movie file of the author dancing with the PHANToM [movetitles.mpg]. The movie displays the move numbers and descriptions as the randomly sequenced moves are danced.

The dance moves last eight beats and are performed to a soundtrack of a well-known song at 120 beats per minute. Each move takes four seconds. Perceptuomotor delays are known to be below 350 ms, so this dance occurs slowly enough that it is possible for the subject to be replaying an open-loop motion program through most of the duration of each move. The choice of upcoming move is not signalled in any way before the new move begins. The timing of a move transition, though, is evident from the soundtrack. For about the first 100 to 350 ms after a move transition, the subject may be doing one of two things: playing a motion program without knowing whether it corresponds to the correct movement; or reacting through the prior settings of impedance parameters on his limb. If it is the latter, the subject can only cognitively react to the sensed input after a minimum long-loop delay. That is, immediately after a move transition the subject can have no valid prediction of which move is being danced.

The random move transitions are a realization of a Markov chain connecting the moves. The transition matrix that generated our PHANToM leader's dance is

$$P = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix} \tag{5.2}$$

This transition matrix does not allow move repetitions, which helps subjects recall that move transitions are scheduled to happen every 8 beats. Every move is equally likely. Also, at the move transition points, the leader is equally likely to reverse direction or to continue

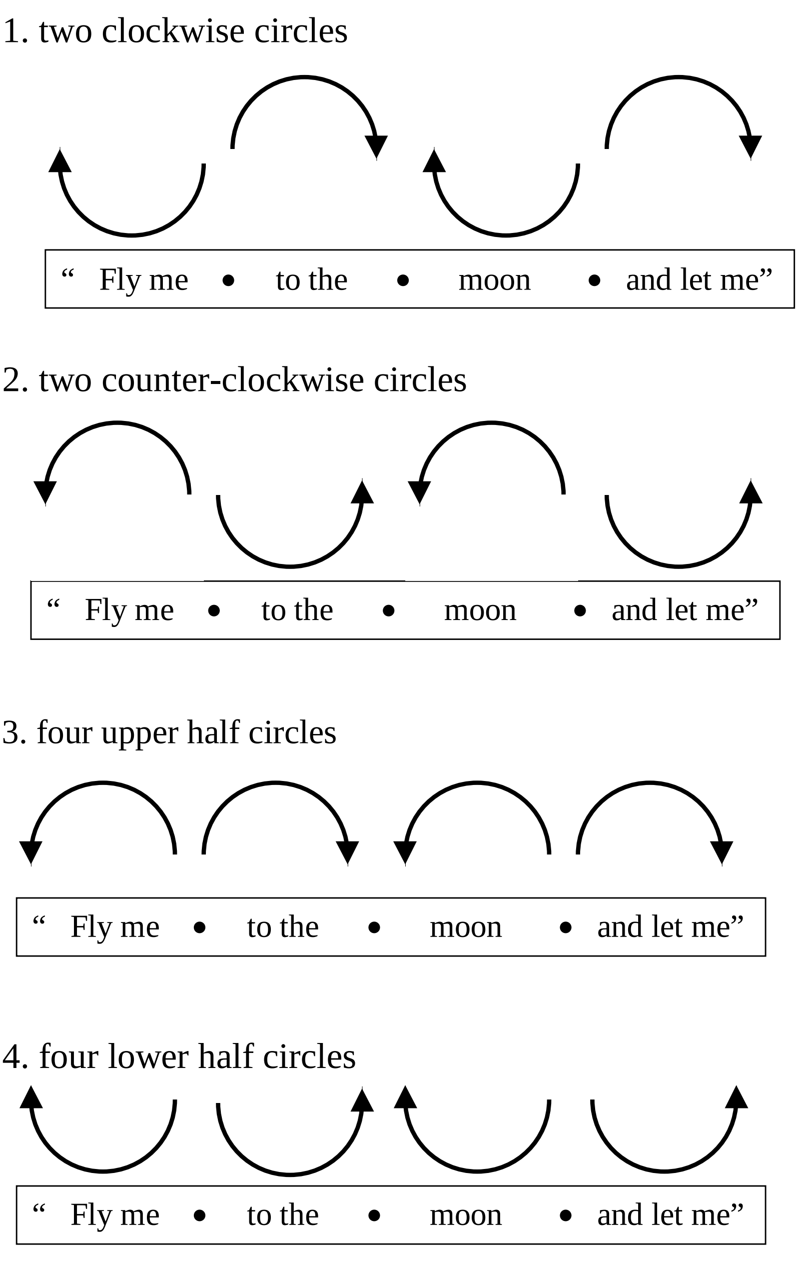

Figure 5-3: Dance moves

in the same trajectory around a circle.

There are three direction reversals within moves 3 and 4, and the very first direction reversal within each move might be predictable, but only if the subject has learned the structure of the transition matrix P. The second and third direction reversals within moves 3 and 4 should definitely be predictable if the subject has learned moves 3 and 4, even absent an understanding of the transition matrix P. In Section 5.5.1, the within move reversals discussed are only the definitely predictable ones. In Section 5.5.2, the question of whether subjects learn the transition matrix is discussed with reference to both definitely predictable and perhaps predictable reversals.

5.4.2 Protocol

The subjects for this experiment were not experienced partner dancers. Perhaps experienced dancers would perform better than the actual subjects did at this dance task, but we aim to apply this novel human-robot interaction idea to more than a small population of trained dancers.

Five subjects volunteered, four male and one female, all of whom were graduate students and researchers familiar with the PHANToM and haptic feedback for virtual environments. The subjects were instructed to try to follow the PHANToM's lead. Subjects were told that the dance was synchronized with the music and consisted of four moves: two circles in each direction, or four half-circles on the upper part of a circle or the lower part of a circle. The subjects first trained for 60 seconds each on the four moves from Figure 5-3. The subjects then trained for 90 seconds each on four two-move fixed sequences, such as two clockwise circles followed by four lower half circles.

The subjects then followed randomly generated move sequences lasting 120 seconds. Participants expressed varying levels of entertainment and frustration in trying to follow the moves.

5.5 Metrics and Results

In pursuit tracking experiments with pseudorandom smooth inputs, researchers have used metrics such as time to task completion, where the task is to move a pointer to within a certain tolerance of zero for 20 seconds, or total mean squared error over the task [34]. For the dance sequence following task, these metrics are inappropriate. Mean squared tracking error sums two sources of error that should be separated: move confusion error and tracking error for executing a move in the case that the move type is known. Another possible metric is time to task completion, where the task is to infer the correct move transition. It is not clear how to calculate the time to task completion metric for our dance, although automatic segmentation and classification methods such as hidden Markov models might be applied to the continuous trajectories. The relevant metrics chosen for the current experiment are time delay in reversing direction, position errors, and variance in position traces.

5.5.1 Reversal Delays

Both at move transitions and within moves, sudden direction reversals are led. Because the transitions between one move and the next move are random, subjects have no way to precognitively or open-loop control the direction reversals at move transitions. However, a direction reversal within a move should be wholly predictable if subjects learn the moves as trajectories. We define the time between reference position reversal and actual y direction (up-down) reversal as the time to task completion. This metric is called *delay* within this section.

The Mann-Whitney-Wilcoxon rank sum test [116] was applied to determine whether the move transition delays and the within move delays could have been sampled from the same distribution. The p-values for the rank sum test were calculated separately for each individual, and the overall p-value for the hypothesis that the distributions were the same was $p = .029$. We conclude that delays at within move reversals are distributed differently than delays at move transitions, so the hypothesis 3 is confirmed.

In Figure 5-4, the distribution (median, interquartile range box, and outliers) of all subjects' move transition and within move direction reversal delays is shown. There are 180 data points for each category included, for 360 total observations. A direction reversal between moves will happen at only about half of these move transitions and is thus unpredictable. The data demonstrate that shorter delays are achievable if the reversal was predicted. Subjects can predict the within move reversals only if they have recognized the current move from their vocabulary of dance moves. Note that the interquartile ranges (IQRs) for these two categories, from the 25th percentile to the 75th percentile, do not overlap.

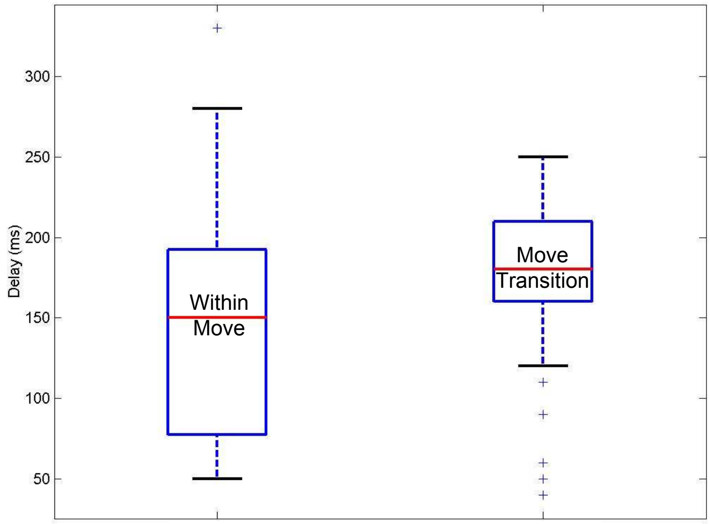

Figure 5-4: Interquartile range boxes, medians, whiskers and outliers of delay. MT, move transitions, are unpredictable direction reversals, while WM, within move direction reversals, should be predictable.

The separation of the IQR boxes indicate recognition of the dance moves. Subjects often predicted and planned for the direction reversal but in some instances reverted to long-loop feedback response, even for within-move reversals, perhaps because they were still novices at the dance. There may also be some artificially low delays included in the move transition delays, reflecting places where a subject erroneously assumed he could predict a direction reversal at a move transition and was, by chance, correct.

The reference velocity jumps at direction reversals, and inertia prevents exact tracking of these sudden velocity changes, even if they are perfectly predicted. The inertia-mandated delay is not known for this setting, though it could be estimated by a multi-joint dynamic model. Also, all of the direction reversals, both within move and at move transitions, are predictable in a sense because they all fall on even beats of the music. If direction reversals can be prepared for, say, by trying to reduce the hand's effective mass, then subjects could prepare for even unpredictable direction reversals (where reversals and non-reversals are equally likely).

Young noted that the most rapid adaptation to sudden control system changes is achieved when subjects know both the pre-change and post-change inputs well and where the transitions are accompanied by an audible signal [143]. That is, even the task at the move transitions should allow rapid response, so the separation observed between move transition and within move delays is significant.

5.5.2 Position errors

If a human follower believed the move to come would be move 1 but in fact it was move 2, the resulting error should be distinguished from simple tracking error. Move misclassifications were clearly recognizable in our data. Examples of a desired response and an obvious misclassification in this experiment's data are shown in Figures 5-5 and 5-6. The data in these figures are from the last reversal within vocabulary element 3, the move with four upper half-circles.

The Figures 5-5 and 5-6 also illustrate that the position traces generally did not track nominal position well. Remember that no visual display of the desired position or the position error was ever given to the subjects. Perhaps because our robot leader penalized velocity error more strongly than position error, the shape of the moves the subjects performed was generally accurate but the placement with respect to the nominal position varied by amounts on the order of a few centimeters.

Defining an error at any reversal as an error in phase with square larger than twice the mean squared error, we counted position errors in the data. Unexpectedly, the frequency of position errors was greater within moves, where the reversals should have been predictable, than at move transitions, where reversals were unpredictable. Table 5.1 shows the average of the subjects' rates of position errors, by trial number and by whether the reversals were predictable. See Section 5.4.1 for an explanation of the difference between reversals that are predictable and those that may be predictable.

Subjects showed learning effects in that error frequency decreased in later trials. The reversals that were not predictable had the lowest error rate, while reversals that may be

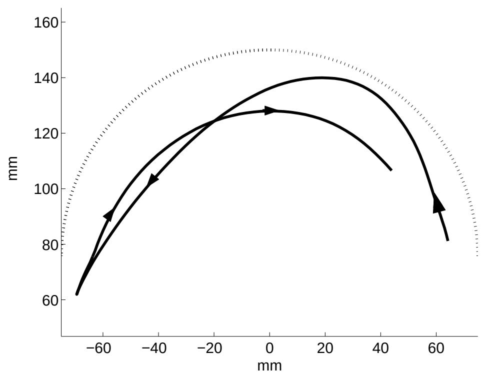

Figure 5-5: Desired reversal (solid line shows position trace, dotted line shows reference position trace)

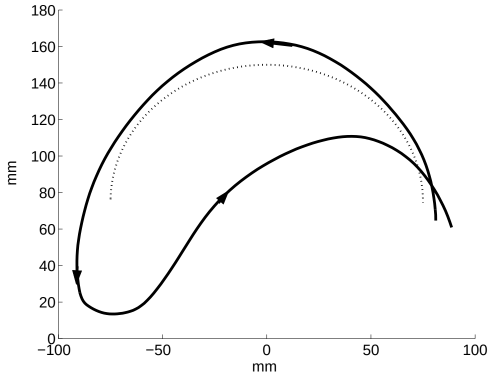

Figure 5-6: Anticipated reversal mistakenly (solid line shows position trace, dotted line shows reference position trace)

predictable and reversals that are predictable had significantly higher error rates. Subjects may have made errors more frequently within a known move because they had a high degree of confidence about the movement. Conversely, they may have made fewer errors when they were aware that they could not predict whether a reversal would occur. Subjects hedged their bets at transitions between moves because they knew they could not predict the transition. Verner et al. found a similar "confidence increases errors" effect in a different task. When gripper-only force feedback was provided in a gripping and manipulation task, errors rose dramatically relative to both the no force feedback and the full force feedback conditions [136]. While it seems that making more errors is a disadvantage of vocabulary-based haptic interaction, the large difference between error rates within moves and at transitions does indicate that subjects altered their strategies once they recognized a move.

Considering the first two trials as practice and the last four trials as valid data, the average rate of position errors by subject appears in Table 5.2. These data are very hard to interpret. All subjects except subject E made no errors at the unpredictable move

Table 5.1: Average large error frequency by trial and predictability

	Not Predictable	May Be Predictable	Predictable
Trial 1	16%	45%	42%
Trial 2	2%	28%	15%
Trial 3	0%	20%	8%
Trial 4	0%	17%	15%
Trial 5	2%	21%	5%
Trial 6	3%	8%	11%

Table 5.2: Average position error frequency by subject and predictability

	Not Predictable	May Be Predictable	Predictable
Subject A	0%	53%	8%
Subject B	0%	4%	1%
Subject C	0%	19%	12%
Subject D	0%	2%	7%
Subject E	5%	6%	20%

transitions. Subjects A and B made more errors at the "may be" predictable reversals than at the predictable reversals, but subjects D and E made more errors at predictable reversals than at the "may be" predictable reversals. This could indicate that subjects A and B were beginning to learn the transition matrix P, which would allow them to predict not only the second and third reversals of moves 3 and 4, but also the very first reversal within moves 3 and 4. Subjects D and E, conversely, performed about the same on these may be predictable reversals as they did on the unpredictable reversals, so these subjects only felt confident about the definitely predictable reversals.

5.5.3 Variance

If subjects are using precognitive tracking in predictable reversals, but can not use precognitive tracking in unpredictable reversals, one might expect to see greater variance in performance at unpredictable reversals. The move density plots in Figures 5-7 through 5-10 each show 20 recorded 1-second intervals after reversals at the right side of the figures, with brighter density where movements coincide. The individual in Figures 5-7 and 5-8 is also the individual with the highest mean phase errors of the five participants, and the distinction between the two plots is obvious to the eye. The individual in Figures 5-9 and 5-10 had smaller mean phase error and less obvious distinction between the two plots. The remainder of the individual variance plot pairs were between these extremes.

These plots were created by gridding the observed x-y positions of recorded 1-second post-reversal intervals with $z = 1$. The falloff functions from the x-y positions were generated by interpolation with zeroes ($z = 0$) randomly placed elsewhere across the grid. The sum of 20 such grids is shown in each figure, with brightness scaled to maximize contrast. These figures may also be explored quantitatively. For instance, the average z value, excluding black grid points, across Figure 5-8 is 7.5 and the average z value, excluding black

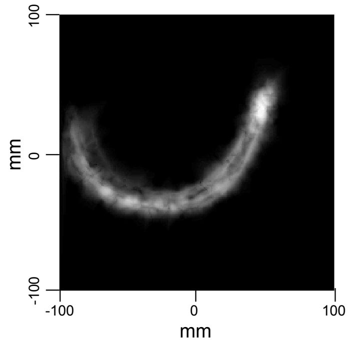

Figure 5-7: A's predictable reversals

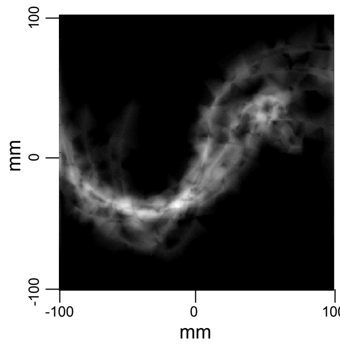

Figure 5-8: A's unpredictable reversals

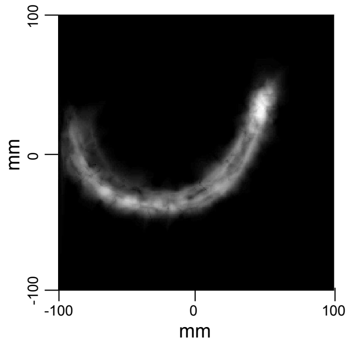

Figure 5-9: B's predictable reversals

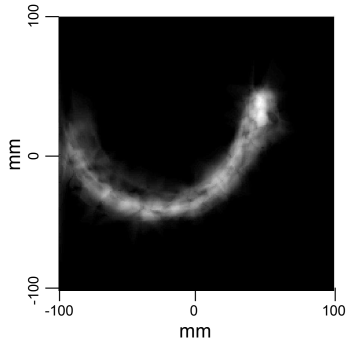

Figure 5-10: B's unpredictable reversals

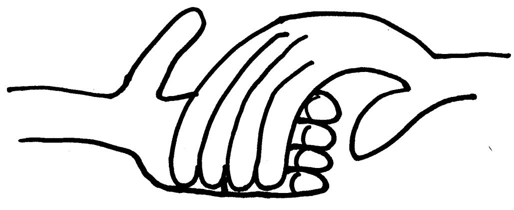

Figure 5-11: Connection: contact without grasping

grid points, across Figure 5-7 is 11.1. This may be interpreted that the average position trace is near to 7.5 other traces in the first case and near 11.1 other position traces in the second case. Grid points at which $z = 0$ are excluded from this calculation because $z = 0$ points would lower the mean if more empty area were gridded, but in fact only the parts of the grid containing traces ($z > 0$) are of interest.

Subject B, whose traces are shown in Figures 5-9 and 5-10, clearly had a different and flatter half-circle in mind than did the former individual in Figures 5-7 and 5-8. This intra-subject variation with relative inter-subject consistency, noted by the authors between all of the subjects, further suggests that individuals respond to an internal model of the trajectory in this dance task. Each subject apparently traced the shape his internal model calls a half-circle.

5.5.4 Connection in z direction

The k_g term increasing the feedback gains is intended to mimic the feel of connection in dance. In partner dancing, the partners' hands remain in contact throughout their motion, even though grasping is not permitted, because each partner imposes a force on the other's hand in opposite directions (see Figure 5-11). An experimental study of a similar type of non-grasping human movement, picking up and moving a box by pressing with an open palm on either side of it, is detailed in Zefran [144], and explored for cooperating robots in Desai [30]. One common connection exercise for partner dancers is for each of them to hold one end of a piece of cloth, and to keep it taut throughout their dance so that the follower can feel small hand movements of the leader which would be imperceptible if the cloth were not taut.

The parallel to dance connection in the PHANToM game is the subject pulling back or pushing in the z direction away from the target plane. The k_g term in (5.1) was intended to encourage subjects to consistently pull away from the target plane in order to get stronger directional signals.

However, no consistent use of the z plane mechanism to increase the force feedback through by the factor k_g was observed. Instead, the z value oscillated with small magnitude at frequencies unrelated to the x and y plane error and position. Subjects did not regulate z position actively, perhaps because grasping the PHANToM stylus is permitted. Subjects did not, therefore, need dance-type connection to maintain contact with the PHANToM. If grasping the stylus had been impossible, subjects would have had to pull or push in the z direction to maintain contact. No participants noticed the small oscillations in z.

5.6 Conclusions

Subjects dancing with the PHANTOM clearly altered their following strategy once they recognized the current move. Subjects reverted to longer-delay, closed-loop or pursuit tracking methods during move transitions at which prediction was not possible. The 75th percentile of the within move delays was lower than the 25th percentile of the move transition delays, indicating that in most cases the within move reversals were predictable because the move had been recognized. Position traces from the experiment showed large intra-subject variability contrasted with strong inter-subject consistency. This suggests that individual conceptions about the circular shape of the move varied, and those conceptions shaped the subjects' interaction with the robot. Responses to unpredictable direction reversals were more variable than responses to predictable direction reversals.

Human-robot interaction designers should exploit the dancing metaphor, because people already demonstrate to great effect their natural skill at haptic coordination in dance. A successful dancing couple is in fact leading and following using a vocabulary of known moves sequenced unpredictably and a musical soundtrack to align timing. A robot facilitated the human experiment in the current work, but the numerous applications of partner dance to human-robot interaction are largely unexplored.

Chapter 6

Haptic collaboration in cyclical Fitts' tasks

Chapter 5 defined haptic collaboration as cooperation between two agents who are physically in contact, whether directly, through a mechanical apparatus, or via a teleoperation system. Dancing Lindy Hop depends critically on the two partners' assisting each other through the moves. The essential move of Lindy Hop is the swingout or whip, which requires the partners to travel from open position to closed position, rotate 360° around each other, and travel back to open position within 8 beats, which is less than 2 seconds at high tempos. This move and many others simply can not be performed by one person without the momentum imparted by the other. Lindy Hop works because the partners exchange energy through their shared dynamics, and so the communication between leaders and followers about what move comes next is not the whole story. The communication and coordination between the partners could not be replaced by pure information transmission, because one partner extends the other's control envelope by adding forces over that same haptic communication channel. In this chapter we focus on a haptic collaborative task which is similar to dancing in that communication happens via forces imposed on a lever the partners share, and in which partners do affect each other's dynamics to extend each other's capabilities.

Understanding how humans assist each other in haptic interaction teams could lead to improved robotic aids to solo human dextrous manipulation. Inspired by experiments reported in Reed et al. [113], which suggested two-person haptically interacting teams could achieve a lower movement time (MT) than individuals for discrete aiming movements of specified accuracy, we report that two-person teams (dyads) can also achieve lower MT for cyclical, continuous aiming movements. We propose a model, called *endpoint compromise*, for how the intended endpoints of both subjects' motion combine during haptic interaction; it predicts a ratio of $\sqrt{2}$ between slopes of MT fits for individuals and dyads. This slope ratio prediction is supported by our data.

6.1 Introduction

A previous study demonstrated the superior performance of two-person haptically interacting teams in a one-dimensional discrete aiming task, also referred to here as a Fitts' task [113]. A discrete aiming task asks a person or a team to move a pointer to a target as quickly as possible with the constraint that the pointer must stop inside the target for some minimum time. In the present experiments, we expand this result to the cyclical aiming domain. Cyclical aiming movements do not require the pointer to come to rest in the target. Instead the pointer should alternate between two targets as quickly as possible without overshooting or undershooting. The task is *not* continuous tracking. The best strategy for low-difficulty cyclical aiming movements is to move in a smooth sinusoidal pattern. Because the pattern of muscle activation required should be substantially different between discrete and cyclical tasks, better performance for dyads in discrete tasks does not imply better performance for dyads in cyclical tasks.

Also, cyclical aiming requires timing coordination between the participants that is not required in discrete aiming. In a discrete aiming task, both people get a visual signal to end one aim and start the next. In cyclical aiming, there is no signal of when to finish aiming at one target and start aiming at the other, except the signals the subjects give to each other by pushing on the device that they share.

6.2 Background

Fitts' law predicts that when humans perform minimum-time aiming tasks of distance D with accuracy specified by the target's width W, the movement time (MT) achieved is

$$MT = A \log_2(\frac{2D}{W}) \tag{6.1}$$

, where the term $\log_2(\frac{2D}{W})$ is referred to as the index of difficulty (ID). Higher difficulty, more accurate movements require longer aiming time. Fitts' law is quite general, in that it has been shown to hold for single-joint and multiple-joint movements, on many scales, across differing pointing devices, and across differing feedback conditions [53]. Fitts' law and its variants have been extensively investigated over the past half-century [61]. This trade-off between speed and accuracy in human motion is an ideal testing ground for experiments elucidating both the limits of human performance and the means by which human performance might be improved by cooperation with other individuals or with well-designed haptic devices.

Discrete Fitts' task performance can be improved by applying a cubic centering force to a joystick used to capture the target [115]. A cubic centering force increases greatly when the joystick is very near the target and tends to push the user back out of the target in the direction from which he entered, making it a virtual fixture. A virtual fixture functions like a ruler, allowing a person substitute what is easy, to hold a certain minimum force against a barrier, for what is difficult, to visually supervise and correct his drawing back to a straight line [120]. Repperger et al. used a cubic centering force to help users accomplish a very simple task [115]. Perhaps humans act as flexible virtual fixtures for each other; this chapter both demonstrates that they enhance each other's performance and suggests a mechanism for how that fixture functions.

6.2.1 Discrete versus cyclical

Fitts first formulated his law for the cyclical case, in which a person is asked to aim successively at each of a pair of targets as rapidly as possible. In the alternative, discrete case, a person is asked to aim at and come to a stop within a given target.

Cyclical Fitts' tasks display an interesting phase shift [53]. Easy cyclical aiming has a harmonic or sinusoidal character, with the maximum acceleration corresponding to the extreme point of each movement. However, as the difficulty of the task increases, cyclical aiming comes to resemble discrete aiming in that each aim comes nearly to a full stop before the following aim begins. Discretization of the movements begins at an index of difficulty between 4 and 5. Guiard has demonstrated that, physically rather than informationally, sinusoidal motion gives cyclical aiming an MT advantage over discrete aiming [53]. This is because sinusoidal motion permits storage and re-use of the kinetic energy a human has generated. Lower movement times are possible for cyclical tasks than for discrete tasks, but only at low difficulty when harmonic motion allows the subject to recycle energy. At higher difficulty, when harmonic motion can not be made accurate enough to hit the target, cyclical aiming has slower movement times than discrete aiming.

Note that harmonic motion has a distinct rhythm. Dyads especially might need to use rhythmic toeholds to enable collaboration, because if each partner started his next aim at an unpredictable time, each would have difficulty predicting the effect of his force input on the device's position. This coordination issue does not arise in the discrete context because in discrete aiming, the start time is displayed to both individuals. Because the present experiments were intended to explore the space of cyclical, harmonic motion in dyads, the index of difficulty was deliberately kept low. IDs of 2.5 to 4.5 were tested in increments of 0.5.

6.2.2 Two-person Fitts' tasks

Mottet et al. first explored two-person, or dyad, cooperative Fitts' tasks [100]. In their design, cooperation was purely informational: one person controlled the motion of the pointer, and another person controlled the motion of the target. Neither could feel the motion of the other; they could only watch the motion of both the cursor and the target on a display. The investigators compared the performance of a dyad, with each subject moving either the target or the pointer, to the performance of a solo subject moving the target with one hand and the pointer with his other hand. Dyads performed the two-handed motion faster than individuals did. This finding was attributed to an information-processing cost to the individual of anti-phase coordination as the individual's two hands moved in opposite directions. However, dyads did not have any advantage over solo subjects who moved only the pointer, with just one hand.

Reed et al. had two subjects move a single pointer into a target region and stop the pointer's motion there [113]. Each of the subjects had a handle to turn the same physical crank, enabling haptic interaction. Dyads performed the movement faster than individuals could. Specifically, the limited data in this experiment suggested that the slope of the Fitts' law curve was the same for dyads and individuals. The dyads aimed an average of 140 milliseconds faster than did individuals. The haptic cooperation speedup was not conclusively explained, though the experimenters speculated that triphasic bursts of effort (agonist for takeoff, then antagonist and agonist for stopping [79]) could be more closely timed if different individuals took responsibility for different phases. Since the pattern of muscle activation for cyclical aiming tasks differs from the sketch above, we experimented with haptically coupled cyclical aiming tasks.

6.2.3 Alternatives to Fitts' model

Fitts' account of the relation

$$MT = A + B\log_2\left(\frac{2D}{W}\right) \tag{6.2}$$

that he discovered in his data was based on Claude Shannon's information theory. The greater the information needed to specify exactly where to stop, Fitts reasoned, the longer the time a human would take to generate that movement. Fitts' law could not explain the variability of movement endpoints, and many other models have been proposed for the speed-accuracy tradeoff [97, 98].

Schmidt et al. found a very different shape for the speed-accuracy tradeoff when subjects were asked to tap alternately across a given distance at a given tempo [121]. Schmidt measured the velocity $\frac{D}{MT}$ and the standard deviation σ of the movements' endpoints, and the data suggested that the latter was proportional to the former: $\sigma \propto \frac{D}{MT}$. That is, a movement's endpoint was approximately normally distributed, and its variance increased with the velocity of the movement. Interpreting the standard deviation of the movements' endpoints as the effective target width W_e, in Schmidt's task there is an approximate linear tradeoff

$$MT = A + B\left(\frac{D}{W_e}\right) \tag{6.3}$$

instead of a logarithmic one.

Our data were best explained by (6.3), using the actual target width, which yields the expression

$$MT = A + B\left(\frac{D}{W}\right) \tag{6.4}$$

.

6.3 Adapting Schmidt's law for dyads

Schmidt's law can make no predictions concerning dyadic performance without a model for how individual movement variabilities would combine during haptic interaction. The individuals are holding the same device, so are constrained to move at the same velocity and share the same movement endpoint. A simple and appealing model is *endpoint compromise*. With the *endpoint compromise* assumption, the individuals move to capture two independently selected movement endpoints, with the result that the pointer stops at the average of the endpoints. We do not offer a physical justification for this assumption, we do hope to address that issue in future work. If σ_d is the standard deviation of the average of two endpoints each with standard deviation σ_s, elementary calculations give

$$\sigma_d = \frac{1}{\sqrt{2}}\sigma_s \tag{6.5}$$

For fixed target width W_0, the individual or dyad is asked to adjust movement time until the probability that the movement endpoint lands outside the target width is less than 5%, but not to move too slowly or aim more accurately than necessary. Lower variability should result in faster movement. Percentiles of the movement endpoint's (normal) distribution scale with standard deviation. Then, for a fixed target width W_0 the solo and dyad movement times MT_s and MT_d should approximately satisfy $MT_d = \frac{1}{\sqrt{2}}MT_s$. *Endpoint compromise* would predict that

$$B_d = \frac{1}{\sqrt{2}}B_s \tag{6.6}$$

where B_d is the slope of (6.4) for MT_d, and the corresponding definition is made for B_s.

6.4 Experiment Protocol and Equipment

Five subjects participated in the experiment; two males and three females, all right-handed except one female was left-handed. Each subject used his dominant hand for all parts of the experiment. The subjects were not paid and had either no or little prior exposure to haptic devices.

Each subject performed a one-dimensional cyclical Fitts' task, aiming alternately at two targets. They used a standard computer driving wheel fixed to a desk, with a 4 ft long wooden dowel attached to create a lever, as pictured in Figure 6-1. The movements were

Figure 6-1: Aiming device; one subject stood on either side of the wheel

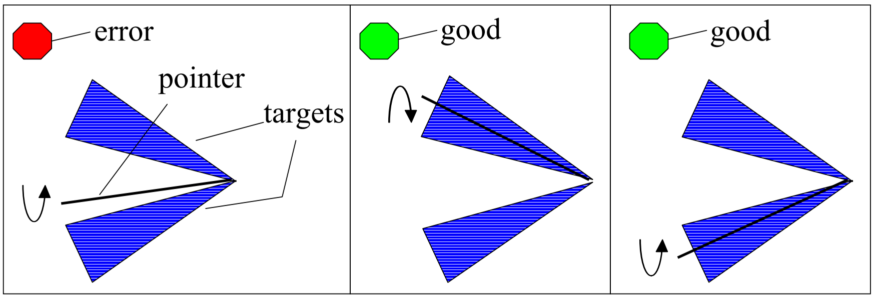

Figure 6-2: Pointer should reverse direction while inside a target

approximately in a plane parallel to the wall behind the desk, so that the subjects aimed alternately up and down at the targets. The two targets were displayed on the computer's screen as sectors of a circle, with the current position marked as a pointer. The display always showed both stationary targets. An indicator in the upper left corner of the screen, colored green or red, corresponded to success or failure of the previous aim. A schematic of the task appears in Fig. 6-2.

Five levels of difficulty were tested. The two targets were fixed at $D = 30°$ apart, and the target widths used were $W = 10.6°, 7.5°, 5.3°, 3.75°$, and $2.65°$, with index of difficulty (ID), defined as $\log_2(\frac{2D}{W})$, of $2.5, 3.0, 3.5, 4.0$, and 4.5. All of the solo and dyad trials were conducted on the same day.

Solo trials

The solo experiment had two practice sessions, each session consisting of five blocks of 60 aims for each of the five difficulty conditions. After each block of 60 aims, the error rate for that difficulty level was displayed. Participants were asked to try to keep their error rates below 5%, but to minimize movement time within that constraint. After the practice sessions, data was collected for 120 aims between the stationary targets for each level of difficulty. During the solo experiment, the subjects all used the left side of the apparatus. Each solo trial lasted between twelve and fifteen minutes.

Dyad trials

Following the solo experiments, each participant was paired with every other participant in the experiment, for a total of ten dyads. Each dyad performed the cyclical aiming task together, with one person on either side of the rotating handle. Because the subjects were on opposite sides of the wheel, when one subject was moving the handle down, the other subject had to be moving the handle up. During the dyad tasks, each person had his own targets and pointer display, corresponding to the position of the handle on his side of the apparatus. The targets always corresponded exactly; there was no conflict between the information displayed to each subject. Subjects were assigned an equal number of times to the left side and to the right side of the apparatus, in alternating order.

The dyad experiment had a single practice session consisting of five blocks of 60 aims for each of the five difficulty conditions. After each block, the error rate for that difficulty level was displayed. The instructions to minimize movement time subject to an error rate below 5% were the same as in the solo experiment. After the practice sessions, data was collected for 120 aims between the stationary targets for each level of difficulty. Each dyad trial lasted between seven and ten minutes. Subjects did not complain of fatigue. Although subjects were invited to rest between any of the blocks of the experiments, only one individual and two dyads stopped for a break during the experiment.

6.4.1 Data Analysis

A computer timer with 10 milliseconds of precision measured the movement durations. The extreme point of each movement marked the end of one aim and the beginning of the next aim. If the extreme point of a movement was within the target sector, then the move was successful; otherwise, the movement was marked as an error. Within the 120 aim trials, the first 20 movements were discarded as warm-up. The remaining movement durations were averaged for each solo and dyad trial.

Table 6.1: Fit of MT (ms) to Fitts' law and Schmidt's law

Solo	$\log_2(\frac{2D}{W})$	$\frac{D}{W}$
Slope	428.2	102.1
Intercept	-658.3	189.7
R^2	0.9320	0.9813
Dyad		
Slope	307.1	76.6
Intercept	-496.4	90.9
R^2	0.8109	0.9250

The angular position of the joystick was recorded as the program ran, but the capture rate was very low, only 12 Hertz. Each movement, therefore, might have had as few as 3 or 4 recorded positions. The harmonicity of the overall movement, however, was clearly visible even with the low capture rate.

A linear least-squares fit between ID and MT, along with an R^2 value, was calculated separately for all five solo trials and all ten dyad trials to assess the validity of Fitts' law with respect to our data. The slopes and intercepts for all solo trials and all dyad trials were averaged. Also, a linear least-squares fit between (D/W) and MT was calculated separately for each trial to assess the validity of Schmidt's law with respect to our data. The slopes and intercepts for all solo trials and all dyad trials were averaged. A summary of these results appears in Table 6.1.

6.5 Results and Discussion

The linear least-squares fit to our movement time data was calculated for Fitts' $\log_2\left(\frac{2D}{W}\right)$ and Schmidt's $\frac{D}{W}$. Table 6.1 shows that the fit to the term $\frac{D}{W}$ is better for both the solo and the dyad conditions. The term $\frac{D}{W}$ explains 98% of the data for individuals and 93% of the data for dyads. The results are most similar to those Schmidt obtained in experiments which were both time-constrained and space-constrained, in which people attempted to move an approximate distance at a given tempo. Table 6.2 contains average movement times, in milliseconds, for every subject and dyad.

No instructions were given regarding verbal negotiation, and a small amount of verbal negotiation did occur. All the conversations surrounded the tempo of the dyad's motion: either, "You're slowing me down," or "We can do this faster."

Table 6.2: Average movement times in milliseconds

Solo	A	B	C	D	E	*Means*
Slope(ms/bit)	107	89.2	129	77.6	108	*102.05*
Intercept (ms)	-56	495	138	101	271	*189.71*
R^2	0.99	0.96	0.99	0.99	0.98	*0.9813*

Dyad	B+A	A+C	C+D	D+A	E+D	A+E	E+C	C+B	B+E	D+B	*Means*
Slope(ms/bit)	89.5	88.5	71.8	83	65.1	82.5	101	67.4	72.3	45.1	*76.6*
Intercept (ms)	122	82.9	108	53.4	92.4	57.8	-45	146	119	172	*90.9*
R^2	0.98	0.92	0.86	0.95	0.96	0.99	0.93	0.87	0.9	0.89	*0.92501*

6.5.1 Errors

None of the individuals or dyads was able to achieve the requested 5% error rate in the difficult tasks. The number of errors and corresponding error rates for all trials are shown in Table 6.3. As in nearly every Fitts' law experiment in the literature, higher error rates were observed for higher difficulty levels. The error rates for dyads and individuals are similar except at the two highest difficulty levels tested, where dyads committed more errors than individuals. This higher error rate might contaminate the finding of lower movement time for dyads, because the dyads did not move as accurately as the solo performers. Future experiments could reward accuracy and provide longer practice times to attempt to remove the accuracy gap between solo and dyad performers.

6.5.2 Movement harmonicity

The angle capture rate was too low to allow estimation of angular position derivatives. Because of the low capture rate, measurements of harmonicity based on time profiles of acceleration as in [53] were not available. The angular traces do afford a qualitative look at the harmonicity of the movements of each individual and dyad.

The range of difficulties in the experiment included tasks difficult enough that harmonic motion was no longer possible for individuals. Figure 6-3 shows a typical trace of an individual doing the most difficult task. Although individuals used inharmonic, discretized motion for the ID = 4.5 task, many dyads achieved harmonic motion for the most difficult task. For instance, Figure 6-4 is an example of fairly smooth sinusoidal motion in the most difficult task. Obviously these characterizations of the traces as harmonic or inharmonic are crude, and some traces seemed to contain a mixture of some harmonic reversals and some discretized reversals, as in Figure 6-5. The profiles which were almost perfectly symmetric about the reversals of direction can be classified as harmonic, while the profiles with significantly more observations (shown as dots in the figures here) before the reversals than after the reversals can be classified as inharmonic.

All individual and dyad traces for the most difficult task, ID = 4.5, were viewed and grouped into three categories: harmonic motion, mixed, and non-harmonic or discretized motion. Of the 5 individual traces, 3 were discretized and 2 were mixed. Of the 10 dyad traces, 5 were harmonic, 3 were mixed, and only 2 were discretized. The categorizations

Table 6.3: Error Counts: 6 errors are 5% of 120 movements

Solo	$\frac{D}{W}$	A	B	C	D	E						*Means*
Errors	*2.81*	3	0	0	2	0						*0.83%*
Errors	*4.00*	7	1	4	1	4						*2.83%*
Errors	*5.66*	4	3	10	1	12						*5.00%*
Errors	*8.00*	14	7	11	10	11						*8.83%*
Errors	*11.32*	9	7	12	8	18						*9.00%*

Dyad	$\frac{D}{W}$	B+A	A+C	C+D	D+A	E+D	A+E	E+C	C+B	B+E	D+B	*Means*
Errors	*2.81*	3	0	1	5	0	4	3	0	3	0	*1.58%*
Errors	*4.00*	2	3	0	1	8	4	8	3	8	0	*3.08%*
Errors	*5.66*	12	8	2	2	4	9	9	4	4	1	*4.58%*
Errors	*8.00*	22	13	9	12	14	16	17	18	13	5	*11.58%*
Errors	*11.32*	10	19	14	15	24	20	28	29	15	15	*15.75%*

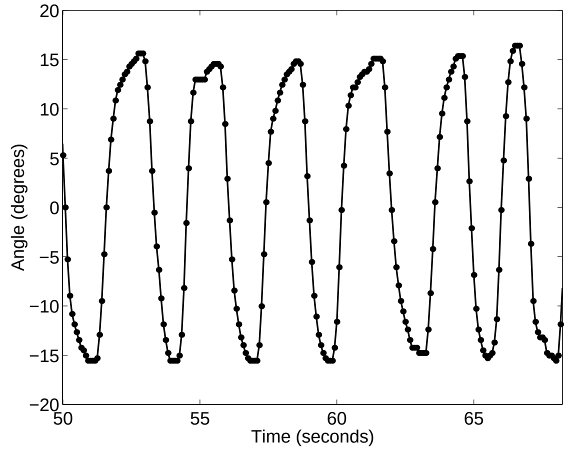

Figure 6-3: Subject B, ID=4.5, inharmonic motion

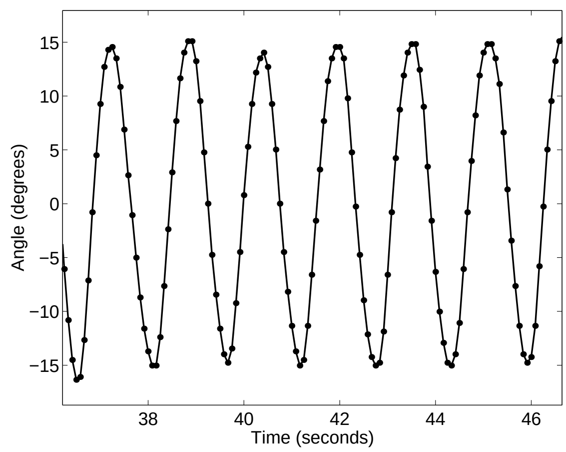

Figure 6-4: Dyad D+B, ID=4.5, harmonic motion

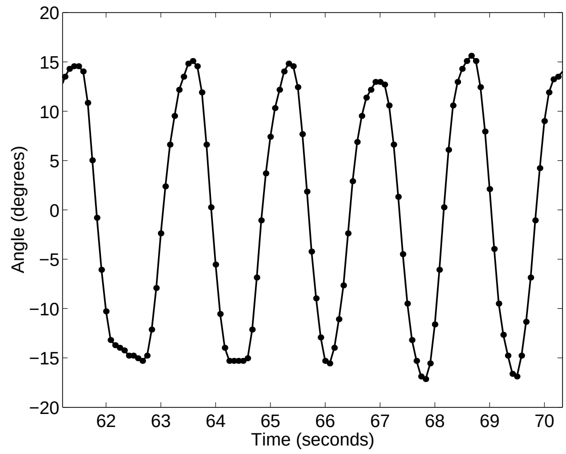

Figure 6-5: Dyad A+B, ID=4, mixed harmonic and discretized

appeared to be unrelated to the error rate. Dyads were sometimes capable of harmonic motion in the hardest task, but no individuals were. This finding is consistent with the observation in Guiard [53] that movement harmonicity is related to high peak velocities. Movement harmonicity only indirectly decreases with the task difficulty because more difficult tasks must be done with lower peak velocities. The peak velocities observed in dyads were usually higher than those observed in solo trials.

6.5.3 Line fits

Dyads' intercepts are roughly 100 milliseconds lower than individuals' intercepts. The average difference between individual and dyad MTs was 99 milliseconds with a 95% confidence interval of (13, 184) milliseconds. The t-distribution (df=18) of this difference was estimated using 20 observations, each a difference between an individual's and one of his dyads's intercept fits. The average MT's for every individual subject appear in Figure 6-6.

Our slope hypothesis, as formulated in Section 6.3, was that the slope of the best fit to MT data for individuals would be larger than the slope of the best fit to MT data for dyads, by a factor of $\sqrt{2} \approx 1.414$. Our data are consistent with this hypothesis. The observed ratio between the solo and dyad slopes was 1.37, with a 95% confidence interval of (1.22, 1.51). The t-distribution (df=18) of this ratio was estimated using 20 observations, each a slope ratio between an individual's and one of his dyads' slope fits. This finding of different slopes is not consistent with the identical slopes reported for the discrete pointing task in [113].

6.5.4 Practice effects and mechanical advantage

Angle and movement time data were collected for all practice phases of the experiment. No significant learning effects were noted in movement harmonicity or error rates. For the solo experiment, movement times increased slightly from the first practice blocks to the final blocks. For the dyad experiment, movement times decreased slightly from the first practice blocks to the final blocks. In both cases, the changes affected primarily the intercepts and not the slopes of the line fits. Since practice tended to make differences between dyad and solo performance larger, it is unlikely practice effects confounded the results. Figure 6-7 shows plots of the average line fits to practice (dashed lines) and actual (solid lines) MTs.

The force required to move the handle did not fatigue the subjects. Still, it is difficult to rule out that the dyads performed better because two people could apply a larger force on the handle to reach a higher peak velocity than was possible for some of the individuals. In Figure 6-6, the lowest individual times were not achieved by the seemingly strongest individual subjects. A refined experimental apparatus might scale by a factor of two the force necessary to move the handle for the dyad case to resolve this question.

6.6 Conclusion

Dyads performed significantly better at a minimum-time cyclical aiming task than individuals. This finding extends the result on two-person improvements for discrete aiming tasks reported by Reed et. al [113]. The slope of the fit to Schmidt's law is lower for dyads than for individuals, and to explain this we propose the *endpoint compromise* hypothesis. Additionally, dyads sometimes maintained a sinusoidal motion at a higher difficulty level than any individual did in these experiments. A physical, dynamical justification for *endpoint compromise* is left to future work.

Accurate aiming, then, is a task which two haptically-interacting people can perform better than one. By investigating a simple reaching task, we probed the lower level communication and coordination mechanisms that might be at work between expert partner dancers. It is conceivable that an endpoint compromise mechanism assists dancers in accurately reaching the desired poses at the required times at fast tempos (see Section 1.1.2).

We enrolled non-dancers in the studies in chapters 5 and 6 to demonstrate that even those without the coordination skills which expert dancers develop can benefit from dance-like haptic interaction with a robot or another person. The haptic-interaction discrete aiming task investigated in prior work [113] lacked the rhythmic coordination challege that the cyclical aiming task imposed. Since dance, like dyadic cyclical aiming, is a rhythmic and communicative task involving shared forces, it would be interesting to repeat this experiment with expert dance partners and compare the dancers' performance to that of non-dancers.

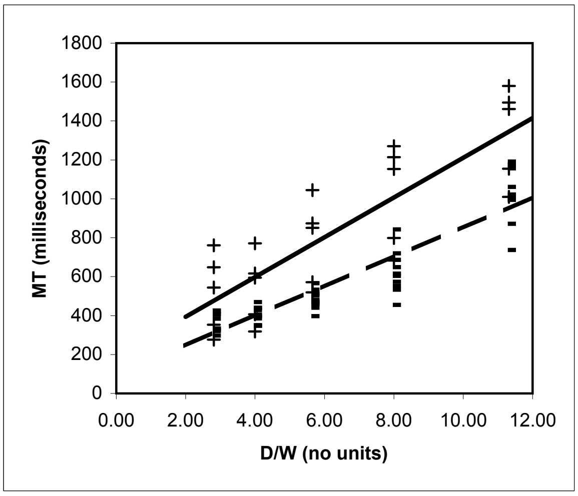

Figure 6-6: Dispersion of average movement times for each subject (plus sign) and each dyad (dash), and averaged line fits for individuals (solid) and dyads (dashed).

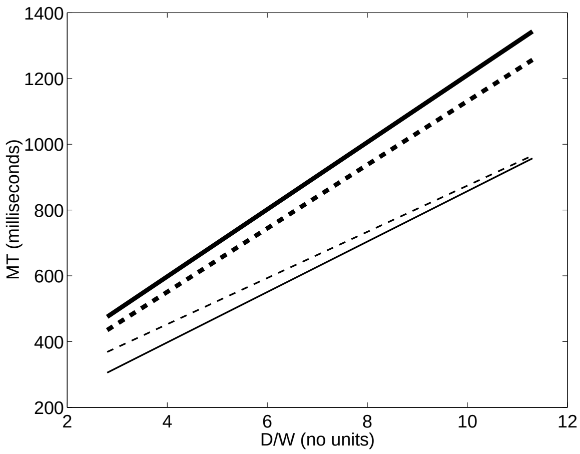

Figure 6-7: With practice, the difference between solo and dyad performance becomes more pronounced. (Solo practice, thick dashed line; solo trials, thick solid line; dyad practice, thin dashed line, dyad trials, thin solid line)

Chapter 7

And the beat goes on

The contributions of the dissertation are two-fold: first, we designed two automated choreographers, and in the process developed useful quantitative models for the high-level choreography decisions that swing dance leaders make. The choreography of a swing dance has a particularly simple structure, since the moves that comprise swing dance and the allowable transitions between moves are actually a finite-state machine. Others have modeled dance and indeed more general human motions as state machines [74, 76], but our state transitions were guided by music interpretation. A novel game theoretic analysis of the interplay between judges, who would prefer to be surprised by the dance choreography, and dancers, who aim for musical appropriateness with each move, in Chapter 4 showed randomness to be an essential feature of pleasing choreography.

Second, we conducted a number of studies addressing the mechanisms of swing dance interaction. This work can be considered elementary scientific inquiry which aimed to characterize touch communication and dynamic interaction between humans and between humans and robots. In Chapter 2 we directly answered the question of whether musical expressiveness is a necessary cue to experienced followers about the intentions of their leaders, at least in Lindy Hop. An experienced follower can detect departures from musical expressiveness, but even when musicality cues are confused she can follow the leader's movements robustly enough that he will not detect any deterioration in the quality of their connection.

Touch communication between partners is necessarily entangled with dynamic interaction. Yet humans performed differently in the two conditions of the experiment in Chapter

5, which specified the dynamics of the robot leader to be exactly the same whether the vocabulary element was predictably or unpredictable. That experiment showed that even non-dancers can make use of a specified motion vocabulary, even in a lead-and-follow situation where elements in that motion vocabulary are led in an unknown order. For the more general task of fast and accurate aiming with two-person haptic interaction considered in Chapter 6, subjects improved each others' solo performance and, interestingly, generated between them a coordinating rhythm that was not provided by a soundtrack.

7.1 Future Work

We have suggested some experimental modifications for exploring alternative explanations of the results we obtained. A number of these were previously described, within the chapters containing each experiment to which these suggestions relate. The automated choreographers we have designed are ready for use in a number of applications beyond the obvious one of musically-tuned dancing figure software widgets.

This research has been able to address only the broadest and most qualitative questions about how leaders and followers swing dance together. A field we would like to pursue, but which we believe may require drastic advances in the modeling of human motor control systems executing whole-body movements [94], is quantitative dynamical analysis of swing dancers and the forces each must exert on each other at a minimum to enable swing dance movements. That is, if it were possible to simulate general multi-person movements of a class that would include most of Lindy Hop, then by searching the parameter space of inputs one could locate the outputs that are similar to movements of experienced and inexperienced dancers. This information would guide dance instructors and judges who currently teach and evaluate on the weight of experience and intuition with little concrete data. Recommendations might even be tuned to an individual dancer's body and strength. It would be exciting to use such a system to classify aerial and acrobatic dance moves according to how well each couple could perform them, so that a couple could concentrate on moves they have the best chance of perfecting.

In fullest generality, we hope that the scientific understanding of partner dances can be advanced on many fronts. A clearer picture of an activity that is one of the wonders of human physical cooperation would be of value to both the dance community and the wider world.

Appendix A: Multimedia additions to the thesis

Audio and video files are included on a CD-ROM that should accompany this thesis. All of the files are noted in the text, near the concepts they illustrate, in square brackets like this: [file.xxx]. Table 1 lists the filenames, file types, and chapter references, along with a short description of what each should contain.

Table 1: Multimedia files

Filename	File Type Description	Chapter
blindfolddance.mpg	Movie (.mpg) blindfolded follower dancing	1.1.4
musicaldancing.mpg	Movie (.mpg) example of musical dance move choices	2.2
interview.mp3	Audio (.mp3) leader interview on musicality experiment	2.5
hijack.mpg	Movie (.mpg) follower hijack reveals song to leader	2.5
autodance.avi	Animated movie (.avi) dance created by automated choreographer	3.5.1
movetitles.mpg	Movie (.mpg) robotic dance movie with text	5.4.1

Bibliography

[1] Jiving with a machine. *Science*, 303:1609, March 2004.

[2] J.J. Abbott and A.M. Okamura. Effects of position quantization and sampling rate on virtual wall passivity. *IEEE Transactions on Robotics*, 2005. In press.

[3] Norman Abramson. *Information theory and coding.* McGraw-Hill, New York, 1963.

[4] R. K. Ahuja and J. B. Orlin. Inverse optimization. *Operations Research*, 49(5):771–783, September/October 2001.

[5] S. Allin, Y. Matsuoka, and R. Klatzky. Measuring just noticeable differences for haptic force feedback: implications for rehabilitation. In *IEEE Haptics*, pages 299–302, March 2002.

[6] Kenji Amaya, Armin Bruderlin, and Tom Calvert. Emotion from motion. In *Graphics Interface*, pages 222–229, May 1996.

[7] Tamer Başar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory.* Society for Industrial and Applied Mathematics, Philadelphia, second edition, 1999.

[8] Margaret Batiuchok. Lindy hop. Master of arts, New York University, May 1988.

[9] Richard Bellman. *Dynamic Programming.* Princeton University Press, Princeton, New Jersey, 1957.

[10] Dmitri P. Bertsekas and John N. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, September 1996.

[11] E. Bizzi and F. A. Mussa-Ivaldi. Geometrical and mechanical issues in movement planning and control. In Michael I. Posner, editor, *The Foundations of Cognitive Science*, pages 769–792. M.I.T. Press, 1989.

[12] Emilio Bizzi, William Chapple, and Neville Hogan. Mechanical properties of muscles: Implications for motor control. *Trends in Neuroscience*, 5(11):395–398, November 1982.

[13] Skippy Blair. *Skippy Blair on Contemporary Social Dance: Disco to Tango and Back.* Golden State Dance Teachers Association.

[14] J.E. Bobrow. Optimal robot motions for physical criteria. *Journal of Robotic Systems*, 18(12), 2001.

[15] A. Borodin and R. El-Yaniv. *Online Computation and Competitive Analysis.* Cambridge University Press, New York, 1998.

[16] Allan Borodin, Morten N. Nielsen, and Charles Rackoff. (Incremental) priority algorithms. In *SODA '02: Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 752–761, Philadelphia, PA, USA, 2002. Society for Industrial and Applied Mathematics.

[17] Matthew Brand and Aaron Hertzmann. Style machines. In *Proceedings of SIGGRAPH*, pages 183–192, New Orleans, Louisiana, 2000. ACM Press / Addison Wesley Longman.

[18] David C. Brogan and Jessica K. Hodgins. Group behaviors for systems with significant dynamics. *Autonomous Robots*, 4:137–153, 1997. Kluwer.

[19] Carolyn Brown, Merce Cunningham, Laura Kuhn, Joseph V. Melillo, Thecla Schiphorst, and David Vaughan. Four key discoveries: Merce Cunningham Dance Company at fifty. *Theater*, 34(2):104–111, July 2004.

[20] Tom Calvert, Lars Wilke, Rhonda Ryman, and Ilene Fox. Applications of computers to dance. *IEEE Computer Graphics and Applications*, 25(2):6–12, March 2005.

[21] Antonio Camurri, Shuji Hashimoto, Matteo Richetti, Andrea Ricci, Kenji Suzuki, Riccardo Trocca, and Gualtiero Volpe. Eyesweb: Toward gesture and affect recognition in interactive dance and music systems. *Computer Music Journal*, 24(1):57–69, 2000.

[22] Antonio Camurri, Ingrid Lagerlf, and Gualtiero Volpe. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, 59(1-2):213–225, July 2003.

[23] Marc Cardle, Loic Barthe, Stephen Brooks, and Peter Robinson. Music-driven motion editing: Local motion transformations guided by music analysis. In *Eurographics Proceedings*, pages 38–44, Leicester, United Kingdom, June 2002.

[24] Diane Chi, Monica Costa, Liwei Zhao, and Norman Badler. The EMOTE model for effort and shape. In *Proceedings of SIGGRAPH*, pages 173–182, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[25] W. J. Conover. *Practical Nonparametric Statistics.* John Wiley and Sons, 2nd edition, 1980.

[26] Microsoft Corporation. Windows Media Player 10. Software, 2004.

[27] R.W. Daniel and P.R. McAree. Fundamental limits of performance for force-reflecting teleoperation. *The International Journal of Robotics Research*, 17(8):811–830, 1998.

[28] George B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, 1963.

[29] Sashka Davis and Russell Impagliazzo. Models of greedy algorithms for graph problems. In *SODA '04: Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 381–390, Philadelphia, PA, USA, 2004. Society for Industrial and Applied Mathematics.

[30] Jaydev P. Desai and Vijay Kumar. Motion planning for cooperating mobile manipulators. *Journal of Robotic Systems*, 16(10):557–579, 1999.

[31] Bruce Randall Donald, James Jennings, and Daniela Rus. Analyzing teams of cooperating mobile robots. In *IEEE International Conference on Robotics and Automation*, pages 1896–1903, 1994.

[32] Gianfranco Doretto, Alessandro Chiuso, Ying Nian Wu, and Stefano Soatto. Dynamic textures. *International Journal of Computer Vision*, 51(2):91–109, 2003.

[33] Fred C. Dyer. The biology of the dance language. *Annual Review of Entomology*, 47:917–949, 2002.

[34] E.C.Poulton. *Tracking Skill and Manual Control*. Academic Press, 1974.

[35] David Eppstein. Setting parameters by example. *SIAM Journal on Computing*, 32(3):643–658, 2003.

[36] J. Alexander Fax and Richard Murray. Information flow and cooperative control of vehicle formations. In *Proceedings of the IFAC World Congress*, Barcelona, Spain, July 2002.

[37] David Feygin, Madeleine Keehner, and Frank Tendick. Haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill. In *IEEE Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 40–47, March 2002.

[38] Tamar Flash. The control of hand equilibrium trajectories in multi-joint arm movements. *Biological Cybernetics*, 57:257–274, 1987.

[39] Tamar Flash and Neville Hogan. The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5(7):1688–1703, July 1985.

[40] Ajo Fod, Maja J. Matarić, and Odest Chadwicke Jenkins. Automated derivation of primitives for movement classification. *Autonomous Robots*, 12:39–54, 2002.

[41] Robert Fourer, David M. Gay, and Brian W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*, 1993.

[42] Tony Fragiacomo. MIDIArt and MIDIWorks. Software, 1993. http://www.midiworks.com/ (Accessed April 19, 2005).

[43] Emilio Frazzoli. *Robust hybrid control for autonomous vehicle motion planning*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, June 2001.

[44] Emilio Frazzoli. Maneuver-based motion planning and coordination for multiple unmanned aerial vehicles. In *AIAA Digital Avionics Systems Conference*, 2002.

[45] Sommer Gentry and Eric Feron. Modeling musically meaningful choreography. In *Proceedings of IEEE Systems, Man, Cybernetics Conference*, October 2004.

[46] Sommer Gentry and Eric Feron. Musicality experiments in lead and follow dance. In *Proceedings of IEEE Systems, Man, Cybernetics Conference*, October 2004.

[47] Sommer Gentry and Roderick Murray-Smith. Haptic dancing: human performance at haptic decoding with a vocabulary. In *Proceedings of IEEE Systems, Man, Cybernetics Conference*, October 2003.

[48] Sommer Gentry, Steven Wall, Ian Oakley, and Roderick Murray-Smith. Got rhythm? Haptic-only dancing with a PHANToM. In *Proceedings of Eurohaptics*, Lecture Notes in Computer Science, July 2003.

[49] R. Brent Gillespie, M. Sile O'Modhrain, Philip Tang, David Zaretsky, and Cuong Pham. The virtual teacher. In *Proceedings of the ASME International Mechanical Engineering Congress: Dynamic Systems and Control Division, DSC-Vol. 60*, volume 2, 1998.

[50] Hiroaki Gomi and Mitsuo Kawato. Equilibrium-point control hypothesis examined by measured arm stiffness during multijoint movement. *Science*, 272:117–120, April 1996.

[51] John H. Gossett, Bradley D. Clymer, and Hooshang Hemami. Long and short delay feedback on one-link nonlinear forearm with coactivation. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(9):1317–1327, 1994.

[52] Ann Hutchinson Guest. *Labanotation: the system of analyzing and recording movement*. Routledge, New York, 4th edition, 2005. Illustrated by Doug Anderson.

[53] Yves Guiard. Fitts' law in the discrete vs. cyclical paradigm. *Human Movement Science*, 16:97–131, 1997.

[54] Gary Guthart and John Kenneth Salisbury Jr. The Intuitive$^{\text{tm}}$ telesurgery system: Overview and application. In *Proceedings of the International Conference on Robotics and Automation*, pages 618–621, 2000.

[55] J.G. Hale and F.E. Pollick. Sticky hands interaction with an anthropomorphic robot. In *Workshop on Interactive Robots and Entertainment*, Carnegie Mellon University, Pittsburgh, Pennsylvania, May 2000.

[56] Blake Hannaford and Jee Hwan Ryu. Time-domain passivity control of haptic interfaces. *IEEE Transactions on Robotics and Automation*, 18(1):1–10, February 2002.

[57] R.A. Hess. Human-in-the-loop control. In William Levine, editor, *The Control Handbook*, pages 1497–1505. CRC Press, 1996.

[58] Jessica K. Hodgins, Wayne L. Wooten, David C. Brogan, and James F. O'Brien. Animating human athletics. In *Proceedings of SIGGRAPH*, pages 71–78, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[59] Neville Hogan. Controlling impedance at the man/machine interface. *IEEE Journal of Robotics and Automation*, pages 1626–31, 1989.

[60] Eugene Hsu, Sommer Gentry, and Jovan Popović. Example-based control of human motion. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 69–77. ACM Press, 2004.

[61] Richard J. Jagacinski and John M. Flach. *Control Theory for Humans: Quantitative approaches to modeling performance*. Lawrence Erlbaum Associates, Mahwah, New Jersey, 2003.

[62] Richard J. Jagacinski, John M. Flach, and Richard D. Gilson. Comparison of visual and kinesthetic-tactual displays for compensatory tracking. *IEEE Transactions on Systems, Man, and Cybernetics*, 13(6):1103–1112, 1983.

[63] Richard J. Jagacinski and Sehchang Hah. Progression-regression effects in tracking repeated patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1):77–88, 1988.

[64] Laci Jalics, Hooshang Hemami, Bradley Clymer, and Alan Groff. Rocking, tapping, and stepping: a prelude to dance. *Autonomous Robots*, 4:227–242, 1997.

[65] Lynette A. Jones. Visual and haptic feedback in the control of force. *Experimental Brain Research*, 130(2):269–272, January 2000.

[66] T. C. Justus and J. J. Bharucha. Modularity in musical processing: the automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4):1000–1011, 2001.

[67] Daniel Kahneman. Maps of bounded rationality. *The Nobel Prizes 2002*, 2003.

[68] Kanav Kahol, Priyamvada Tripathi, and Sethuraman Panchanathan. Automated gesture segmentation from dance sequences. In *Proceedings of 6th IEEE Conference on Automatic Face and Gesture Recognition*, 2004.

[69] P. Kall and S. W. Wallace. *Stochastic Programming*. John Wiley and Sons, 1994.

[70] A. Karniel and F.A. Mussa-Ivaldi. Does the motor control system use multiple models and context switching to cope with a variable environment? *Experimental Brain Research*, 143(4):520–524, 2002.

[71] H. Kazerooni. Human-robot interaction via the transfer of power and information signals. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2):450–463, 1990.

[72] H. Kazerooni and Tanya J. Snyder. Case study on haptic devices: Human-induced instability in powered hand controllers. *Journal of Guidance, Control, and Dynamics*, 18(1):108–113, February 1995.

[73] Christopher W. Kennedy, Tie Hu, Jaydev P. Desai, Andrew S. Wechsler, and J. Yasha Kresh. A novel approach to robotic cardiac surgery using haptics and vision. *Cardiovascular Engineering*, 2(1):15–22, March 2002.

[74] Tae-hoon Kim, Sang Il Park, and Sung Yong Shin. Rhythmic-motion synthesis based on motion-beat analysis. In *Proceedings of SIGGRAPH*, volume 22, pages 392–401, 2003.

[75] Kazuhiro Kosuge, Tomohiro Hayashi, Yasuhisa Hirata, and Ryosuke Tobiyama. Dance partner robot -Ms DanceR-. In *Proceedings of IEEE Conference on Intelligent Robots and Systems*, volume 4, pages 3459–3464, October 2003.

[76] Lucas Kovar, Michael Gleicher, and Frederic Pighin. Motion graphs. In *Proceedings of SIGGRAPH*, volume 21, New York, NY, USA, July 2002. ACM Press/Addison-Wesley Publishing Co.

[77] Arthur D. Kuo. An optimal control model for analyzing human postural balance. *IEEE Transactions on Biomedical Engineering*, 42(1):87–100, January 1995.

[78] Arthur D. Kuo. The relative roles of feedforward and feedback in the control of rhythmic movements. *Motor Control*, 6:129–145, 2002.

[79] N. Lan and P.E. Crago. Optimal control of antagonistic muscle stiffness during voluntary movements. *Biological Cybernetics*, 71(2):123–135, 1994.

[80] Ning Lan. Analysis of an optimal control model of multi-joint arm movements. *Biological Cybernetics*, 76:107–117, 1997.

[81] T. Längle and H. Wörn. Human-robot cooperation using multi-agent-systems. *Journal of Intelligent and Robotic Systems*, 32:143–159, 2001.

[82] John Lansdown. The computer in choreography. *Computer*, 11(8):19–30, August 1978.

[83] Dale A. Lawrence. Stability and transparency in bilateral teleoperation. *IEEE Transactions on Robotics and Automation*, 9(5):624–637, October 1993.

[84] Kenneth Laws. An analysis of turns in dance. *Dance Research Journal*, 11:12–19, 1978–1979.

[85] Kenneth Laws. *The physics of dance*. Schirmer Books, 1984.

[86] M. Li and A. Okamura. Recognition of operator motions for real-time assistance using virtual fixtures. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems Symposium*, 2003.

[87] Yan Li, Tianshu Wang, and Heung-Yeung Shum. Motion texture: a two-level statistical model for character motion synthesis. In *Proceedings of SIGGRAPH*, volume 21, pages 465–472, New York, NY, USA, July 2002. ACM Press/Addison-Wesley Publishing Co.

[88] W. Lytle. *Driving computer graphics animation from a musical score*, volume 2 of *Scientific Excellence in Supercomputing: The I B M 1990 Contest Prize Papers*. 1991.

[89] W. Lytle. Animusic: a computer animation video album. DVD, 2001.

[90] R.E. Magdaleno, H.R. Jex, and W.A. Johnson. Tracking quasi-predictable displays. In *Fifth annual NASA University Conference on Manual Control*, pages 391–428, 1969.

[91] P. Marayong and A. Okamura. Speed-accuracy characteristics of human-machine cooperative manipulation using virtual fixtures with variable admittance. *Human Factors*, 46(3):518–532, 2004.

[92] Steve G. Massaquoi and Jean-Jacques Slotine. The intermediate cerebellum may function as a wave-variable processor. *Neuroscience Letters*, 215:60–64, 1996.

[93] Thomas H. Massie and J.K. Salisbury. The phantom haptic interface: A device for probing virtual objects. In *Proceedings of the ASME Winter Annual Meeting, Symposium of Haptic Interfaces for Virtual Environment and Teleoperator Systems*, November 1994.

[94] Maja J. Matarić, Victor B. Zordan, and Matthew W. Williamson. Making complex articulated agents dance. *Autonomous Agents and Multi-Agent Systems*, 2:23–43, 1999.

[95] Joseph McIntyre and Emilio Bizzi. Servo hypotheses for the biological control of movement. *Journal of Motor Behavior*, 25(3):193–202, 1993.

[96] Craig R.M. McKenzie. Rational models as theories - not standards - of behavior. *Psychological Review*, 7(9):403–406, September 2003.

[97] David E. Meyer, Sylvan Kornblum, Richard A. Abrams, Charles E. Wright, and J.E. Keith Smith. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological Review*, 95(3):340–370, 1988.

[98] David E. Meyer, Sylvan Kornblum, Richard A. Abrams, Charles E. Wright, and J.E. Keith Smith. Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In M. Jeannerod, editor, *Attention and Performance XIII*, volume 13, pages 173–226, Hillsdale, NJ, 1990. Erlbaum.

[99] L. B. Meyer. On rehearing music. In *Music, the arts, and ideas*, pages 42–53, Chicago, 1967. University of Chicago Press.

[100] Dennis Mottet, Yves Guiard, Thierry Ferrand, and Reinoud J. Bootsma. Two-handed performance of a rhythmical fitts task by individuals and dyads. *Journal of Experimental Psychology: Human Perception and Performance*, 27(6):1275–1286, 2001.

[101] F. Mussa-Ivaldi. Nonlinear force fields: a distributed system of control primitives for representing and learning movements. In *Proceedings of IEEE Symposium on Computational Intelligence in Robotics and Automation*, Monterey, California, 1997.

[102] F. Mussa-Ivaldi, N. Hogan, and E. Bizzi. Neural, mechanical, and geometrical factors subserving arm posture in humans. *Journal of Neuroscience*, 5:2732–2743, 1985.

[103] Darrell Nakawaki, Sangwan Joo, and Fumio Miyazaki. Dynamic modeling approach to gymnastic coaching. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 1069–1076, 1998.

[104] Günter Niemeyer and Jean-Jacques E. Slotine. Telemanipulation with time delays. *International Journal of Robotics Research*, 23(9):873–890, September 2004.

[105] A. Michael Noll. Choreography and computers. *Dance Magazine*, 41(1):43–45, January 1967.

[106] Ian Oakley, Stephen Brewster, and Philip Gray. Haptic gesturing. University of Glasgow technical report, 2001.

[107] Allison M. Okamura. Methods for haptic feedback in teleoperated robot-assisted surgery. *Industrial Robot*, 31(6):499–508, 2004.

[108] Alex Pentland and Andrew Liu. Modeling and prediction of human behavior. *Neural Computation*, 11:229–242, 1999.

[109] Ken Perlin. Real time responsive animation with personality. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):5–15, January 1995.

[110] Katherine Pullen and Christoph Bregler. Motion capture assisted animation: texturing and synthesis. In *Proceedings of SIGGRAPH*, volume 21, pages 501–508, New York, NY, USA, 2002. ACM Press.

[111] Christine C. Raasch, Felix E. Zajac, Baoming Mao, and William S. Levine. Muscle coordination of maximum-speed pedaling. *Journal of Biomechanics*, 30(6):595–602, 1997.

[112] Jens Rasmussen. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, 13(3):257–266, 1983.

[113] Kyle Reed, Michael Peshkin, J. Edward Colgate, and James Patton. Initial studies in human-robot-human interaction: Fitts' law for two people. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 2333–2338, 2004.

[114] R. Reiner, T. Nef, and G. Colombo. Robot-aided neurorehabilitation of the upper extremities. *Medical and Biological Engineering and Computing*, 43:2–10, January 2005.

[115] D.W. Repperger, Chandler A. Phillips, and Tamara L. Chelletter. A study on spatially induced 'virtual force' with an information theoretic investigation of human performance. *IEEE Transactions on Systems, Man, and Cybernetics*, 25(10):1392–1404, October 1995.

[116] John A. Rice. *Mathematical Statistics and Data Analysis*. Wadsworth Publishing Company, Belmont, CA, second edition, 1995.

[117] K. Rohrseitz and J Tautz. Honeybee dance communication: waggle run direction coded in antennal contacts? *Journal of Comparative Physiology*, 184:463–470, 1999.

[118] Charles Rose, Michael F. Cohen, and Bobby Bodenheimer. Verbs and adverbs: Multidimensional motion interpolation. *IEEE Computer Graphics and Applications*, 18(5):32–40, September 1998.

[119] J. Rosen, M. Solazzo, B. Hannaford, and M. Sinanan. Task decomposition of laparoscopic surgery for objective evaluation of surgical residents' learning curve using hidden markov model. *Computer Aided Surgery*, 7, 2002.

[120] Louis B. Rosenberg. Virtual fixtures: perceptual tools for telerobotic manipulation. In *Proceedings of IEEE Virtual Reality Symposium*, pages 76–82, 1993.

[121] R.A. Schmidt, H.N. Zelaznik, B. Hawkins, J.S. Frank, and J.T. Quinn Jr. Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, 86:415–451, 1979.

[122] Arno Schodl, Richard Szeliski, David H. Salesin, and Irfan Essa. Video textures. In *Proceedings of SIGGRAPH*, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[123] S.A. Setiawan, J. Yamaguchi, Sang Ho Hyon, and A. Takanishi. Physical interaction between human and a bipedal humanoid robot. In *Proceedings of IEEE Conference on Robotics and Automation*, volume 1, pages 361–367, 1999.

[124] M.E. Sezer and D.D. Siljak. Decentralized control. In *The Control Handbook*, pages 779–793. IEEE Press, 1996.

[125] Thomas B. Sheridan. *Telerobotics, Automation, and Human Supervisory Control.* MIT Press, 1992.

[126] Thomas B. Sheridan and William R. Ferrell. *Man-machine systems: information, control, and decision models of human performance.* MIT Press, Cambridge, Massachusetts, 1981.

[127] T. Sugar and V. Kumar. Decentralized control of cooperating manipulators. In *Proceedings of the IEEE International Conference on Robotics and Automation*, Leuven, Belgium, May 1998.

[128] R. Taylor, P. Jensen, L. Whitcomb, A. Barnes, R. Kumar, D. Stoianovici, P. Gupta, Z. Wang, and L. Kavoussi. A steady-hand robotic system for microsurgical augmentation. *International Journal of Robotics Research*, 18(12):1201–1210, 1999.

[129] F. Towhidkhah, R. E. Gander, and H. C. Wood. Model predictive impedance control: a model for joint movement. *Journal of Motor Behavior*, 29(3):209–222, 1997.

[130] Aydin Tözeren. *Human Body Dynamics.* Springer, 2000.

[131] Ryan Ulyate and David Bianciardi. The interactive dance club: avoiding chaos in a multi-participant environment. *Computer Music Journal*, 26(3):40–49, 2002.

[132] Y. Uno, M. Kawato, and R. Suzuki. Formation and control of optimal trajectory in human multijoint arm movement. minimum torque-change model. *Biological Cybernetics*, 61(2):89–101, 1989.

[133] Michiel Van de Panne. Control for simulated human and animal motion. *Annual Reviews in Control*, 24:189–199, 2000.

[134] R.Q. van der Linde and P. Lammertse. Haptic Master - a generic force controlled robot for human interaction. *Industrial Robot*, 30(6):515–524, 2003.

[135] D. Del Vecchio, Richard M. Murray, and P. Perona. Decomposition of human motion into dynamics based primitives with application to drawing tasks. *Automatica*, 39(12):2085–2098, 2003.

[136] L. Verner, K. A. Jeung, and A. M. Okamura. The effects of gripping and translational forces on teleoperation. In *Workshop on multi-point interaction in robotics and virtual reality at IEEE International Conference on Robotics and Automation*, New Orleans, Louisiana, April 27 2004. In press as a chapter in the upcoming book *Multi-point Interaction in Robotics and Virtual Reality* (Springer).

[137] Christopher R. Wagner and Robert D. Howe. Mechanisms of performance enhancement with force feedback. In *First Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Pisa, Italy, March 2005.

[138] C.R. Wagner, N. Stylopoulos, and R.D. Howe. The role of force feedback in surgery: analysis of blunt dissection. In *Proceedings of 10th Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems*, pages 68–74, Orlando, FL, March 2002.

[139] Qi Wang, Naoki Saiwaki, and Shogo Noshida. An approach to automatic motion synthesis for 3D spring model. In *Proceedings of IEEE International Confererence on Systems, Man, and Cybernetics*, volume 4, pages 224–229, Tokyo, Japan, 1999.

[140] Robert Wechsler. Computers and art: a dancer's perspective. *IEEE Technology and Society Magazine*, Fall 1997.

[141] Robert Wechsler. O body swayed to music (and vice versa): roles for the computer in dance. *Leonardo*, 30(5), October 1997.

[142] Robert L. Williams, Mayank Srivastava, Robert R. Conatser, Jr., and John N. Howell. Implementation and evaluation of a haptic playback system. *Haptics-e*, 3(3), 2004.

[143] L.R. Young. On adaptive manual control. *IEEE Transactions on Man-Machine Systems*, 10(4):292–331, 1969.

[144] M. Zefran, V. Kumar, G. Garvin, and J. Desai. Experimental investigation of human two-arm manipulation. In *International Conference on Mechanics in Medicine and Biology*, 1996.

[145] Liwei Zhao and Norman I. Badler. Acquiring and validating motion qualities from live limb gestures. *Graphical Models*, 67(1):1–16, January 2005.