Seeking Allies: Modelling how listeners choose their musical friends

Dave Billinge
Department of Creative Technology,
University of Portsmouth,
Buckingham Building, Burnaby Road,
Portsmouth PO1 3AE
Email: dave.billinge@port.ac.uk

Tom Addis,
Department of Computer Science and Software Engineering,
University of Portsmouth,
Buckingham Building, Burnaby Road,
Portsmouth PO1 3AE

Visiting Research Fellow Science Studies Centre,
Department of Psychology, University of Bath
Email: tom.addis@port.ac.uk

Abstract

In this paper we describe in some detail a formal computer model of inferential discourse based on a belief system. The key issue is that a logical model in a computer, based on rational sets, can usefully model a human situation based on irrational sets. The background of this work is explained elsewhere, as is the issue of rational and irrational sets (Billinge and Addis 2004, Stepney et al 2004). The model is based on the Belief System (Addis and Gooding 1999) and it provides a mechanism for choosing queries based on a range of belief. We explain how it provides a way to update the belief based on query results, thus modelling others' experience by inference. We also demonstrate that for the same internal experience, different models can be built for different actors.

1. Introduction

The problem of what information is exchanged between people talking about music arose when we started to investigate the possibility of providing a computer aid to help music planners devise acceptable music programmes (Billinge 2000, Billinge & Addis 2003). In order to create such a computer aid we needed to formalise the way people perceived music and communicated their perceptions to each other. Studies of people attempting to pass this information seemed to fail completely and further, no correlates were discovered between the words used and the music features (e.g. minor chords relating to sadness). It seemed that talking about music had no effective role and yet people do talk and there is whole industry devoted to communicating about the subjective perception of music. Our observations did not make sense and this required us to reconsider our methods.

The essence of the original approach was to take the simple surface observations of communication, such as words, taken under controlled experimental conditions and then apply statistical and linguistic analyses based on simple denotational semantics. Since these analyses failed to produce a result we then considered a deeper approach where we posited a mechanism of inferential semantics. The mechanism was proposed on the observation that communication of music was rich with metaphor as a descriptive aid; metaphors that can be drawn from more explicit and positive domains such as war or nature. Meaning was thus inferred from relationships evoked by these metaphors that can then be applied to describing unobserved and less concrete ideas such as music perception. The problem with metaphors is that they are a culturally based and dynamic trope and it is because of this that language falls firmly into the area of irrational sets with the consequent difficulties. A technical solution invokes the process of tracking meaning through a belief system (Addis et al 2004, Billinge and Addis 2004, Stepney et al 2004).

We required to show that an inferential semantics needed to use metaphor could work and to this end we harnessed two distinct processes. The first was a belief system originally created to show how scientists decide what experiments to perform or with whom to communicate in order to find out which of several possible hypotheses about the world is workable. Workable here means making the world more predictable. The second mechanism was the internal modelling of other people's beliefs derived from conversation. This latter process was bypassed in the original belief system by assuming that the model would be the same as the actual perceptions. These actual perceptions were made accessible within the computer model by allowing the computer actors (agents) to have partial information of another's perceptions. How this might be accomplished with people was not considered until the second mechanism of modelling people's beliefs was designed. These processes were emulated as computer programs to show how they might work in practice. For such complex processes involving many different scenarios it was only through running such computer models that these processes could be tested for coherence.

Constructing a computer program as though it were a theory is not enough. Theories also have to be tested against the world and the effectiveness of a theory can only be assessed in its ability to make successful predictions. Any theory that makes the world predictable is useful and a better theory will improve on this. A theory can also provide a framework in which to design experiments and recognise significant features. Even a useless theory can play this latter role and without designed experiments and puzzling observations a new theory cannot be coherently created. To this end we produced a computer model of music communication not to say 'This is how it is with people' but to say 'This makes the world a less surprising place' (Peirce, in Weiner 1966).

2. The Experiment

The experiment is describe in detail elsewhere (Billinge & Addis 2003, 2004). A summary of the process is that four people were asked to listen to four pieces of classical (but little known) music. Each piece of music lasted about ten to fifteen minutes. Each person was asked to keep abbreviated notes on what they heard and to rate the music on a single linear scale ranging form 1 to 10 in units of 1. Evaluation of two pieces of music could be the same. They were then asked to form a committee to discuss the music with the purpose of recommending that one of the pieces should be included in a concert programme. The discussion was recorded. For completion the committees as units were asked to rank all four pieces of music on a scale 1 to 10.

The final part of the experiment, and as it turned out the most significant, was to ask the individual participants to rank their fellow committee members in terms of whose judgement they would take the most notice of when deciding to go to a concert. This will be referred to as 'ally choice'

The design of the experiment was primarily to explore the use of metaphor during the discussion and the recordings taken during these discussions have yet to be analysed. However, we were interested in the relationship between the music and ranking. We were also interested in how the discussions might influence opinion, hence the individual-before and the group-after rankings. The test of the effectiveness of a person's internal models of others was to be assessed from the individual ranking of a chosen advisor.

The delay in analysing the discussions was done because we wished to view these data in the light of a computer model of inferential semantics. There are arguments to suggest that this is bad practice since we are prejudicing our observations with the computer model. Such prejudice will cause us to observe only that which will support the model and thus our observations will be tainted and in doubt. We reject such an argument because the history of science supports the need for an initial theory (Kuhn 1985), provided such a theory can be tested and has the possibility of being rejected (Popper 1959). Theories are particularly helpful in observing complex situations, such as group discussions, because they do limit what should be observed. The real test of a theory has been discussed above but a theory also plays a further important role by providing a basis for puzzlement and modification. If you don't have a theory then you cannot be puzzled by what you observe. We therefore, required a satisfactory model of discourse that could do the job of limiting what we observed (at least initially) in the conversation.

3. The Computer Model

The process that models the inferential semantics is driven by the belief system (Addis and Gooding 1999, 2004). The belief system, in this case, models each user as a single undefined dimension of values for each piece of music involved in a discussion. The values on the dimension are discrete and ordered. Each value is considered an independent hypothesis that has a level of belief associated with it. The actions open to the belief system in order to update these associated beliefs is limited as to a fixed set of queries that may be addressed to another person. Person in this case is a part of the computer program that is normally referred to as 'Agent' but we will use the more appropriate and original term 'Actor' (Hewitt 1979). The range of queries given provides a basis for the process of *question* and *inference* required to update the model.

We use the initial decisions made by the participants during their evaluation of the music heard in experimental sessions to initiate the model. We then compare the predicted order of ally choice made by each agent in the model as assessed from the agent's perception of others, as drawn from a simulated conversation, with the actual outcomes of the sessions. We plan to modify the model to show from recorded conversations of the participants in our experimental sessions how the patterns of questioning compare with that generated by the model. The model has n actors and m aesthetic objects. The m objects in our experiments are pieces of classical orchestral music to which our actor/participants have a response. Actors have a response-scale for each piece of music representing their own subjective impression. Further, an actor has a separate scale of belief for the response to each piece of music for each other actor as derived from the conversation. In the model the actors can ask questions in turn and can only ask one question per turn. Only the actor asking the question can update its scales of belief from another actor's response.

We make three assumptions: first that each actor assumes that other actors initially have the same perceptions of the heard music and thus the same ratings; second that each modelling actor tends to ask the other actor about the music of which it, the questioner, has the most uncertain belief scale; and finally that the modeller can have no doubt of its *own* experience. In this way an actor tracks the subjective experience of others; an experience that may change over time. We also assume that, as supported by our observation of the experimental work with people, an actor will choose as an ally the fellow actor who is closest in m scale distance. The significance of this result is that it shows subjective[1] experience can be inferred through conversations and we suggest that this may be a major purpose of peoples' discussions.

Finally we describe how we expect to adapt the model to fit in with observations to take into account other factors that decide group decisions. Our model should thus become more able to predict conversational behaviour and final group decisions from knowing individuals' perceptions within this scenario.

4. Experimental results analysis:

Seven experiments were carried out with four people and one with only three people. Where ally-ranking information is important only the seven fully populated experiments were considered.

[1] An evaluation of a potentially shareable event that is accessible only by a single actor and related to that actor's observation and assessment of that event

The basic ordinal and numerical data from these experiments includes:

- Each subject's personal scale for each piece of music.
- The order in which the music was played.
- The evaluation order for the pieces of music agreed by the group after discussion.
- The order of preference for another actor as mentor or ally.

In order to make ordinal comparisons without the problems of individual scale choice we use a relative z-score. This is used differently from the normal use of a z-score since we consider each individual subject as though they had their own personal distribution. Thus each subject's personal scale is normalised according to the following equation.

$$\text{z-score}(i) = (x_i - \mu_i) / \sigma_i$$

where i is a particular actor, x is a value given by that subject for a piece of music, μ_i is the mean of the subject's values and σ_i is the standard deviation of the values adjusted for small sample of four such that $\sigma_i = \sigma / \sqrt{n}$ where n is 4 in this case.

In this way all the scores are normalised such that all their scores are:

- distributed about a common mean of zero
- the spread of their evaluations is made equal.

Thus the only significant information is the ordering. However, since all scales are now normalised the relative ordering (nearer or further from other pieces of music) information can also be compared. Each subject is represented as a *single point* in a four-dimensional music space where each dimension represents one of the pieces of music they are judging.

Having got rid of personal scaling differences we can now see if there are any similarities in choice combinations. This is to see how independent the four dimensions are. We ask the question, "Does knowing a person's first choice make it possible to predict their second choice?" However, from the correlation analysis of pairing pieces of music, no significant correlation is found. This indicates that there is no pattern of common approval between items; i.e. approval of Stravinsky does not imply approval of, say, Stenhammar. (See example in figure 1).

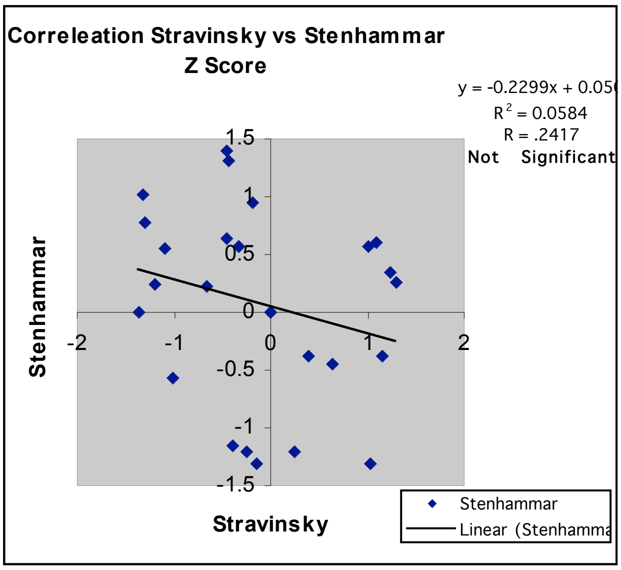

Figure 1: Correlation between two pieces of music

This result was also confirmed using the raw (non z-score normalized) analysis and using optimised principle component analysis (Billinge and Addis 2004). This confirms we can treat music space as set of independent dimensions so that distance calculations in this space conform to normal n-dimensional geometry.

The raw un-normalised values were used to see if there was any correlation between the group scaling and the average of all the individuals' scales. We find that these are positively correlated (see Figure 2) suggesting that participation within the discussions have some influence. Significance r = 0.7584. For 22 experiments an r-value of 0.6524 is better than 0.001 probability that a correlation exists, i.e. there is a 0.1% chance of this happening accidentally. For 24 experiments this would be better, i.e. less likely to occur by chance.

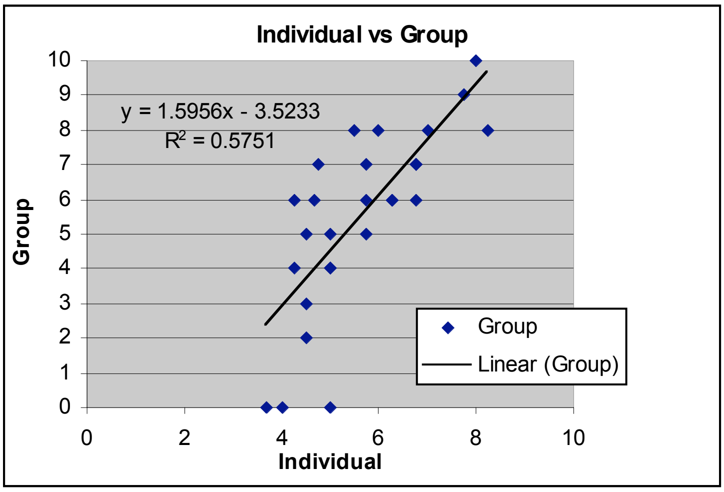

Figure 2: Group scale vs. all individual average

In contrast and as we have noted, previous findings (Billinge & Addis, in George 2005) indicated that no information about music experience or content is exchanged during conversation. A further question is posed "If musical experience is not being exchanged then what influences the group decision?" We speculate that the discussion is thus performing another task. Since it is not concerned with any group experience, could it be concerned with just the exercise of social dominance?

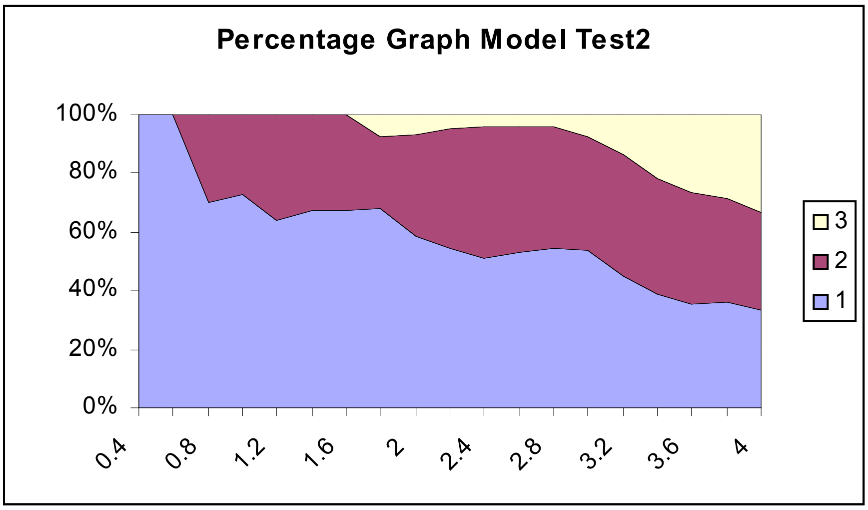

Figure 3: 100% Stacked Area Chart of Evaluative Distance [X-axis] against Likelihood of Choice [Y-axis]

We observe that the discussion proceeds in all seriousness and that there is an outcome in the form of ally choice that was not previously acknowledged. The actor's ally preference order seems to relate to the similarity of the order evaluation of the music between them. Plotting the distance between two subjects in a relative 4-dimensional (music) z-space can assess this order similarity. The hypothesis here is that an actor tends to choose others who have a similar view of the music. For the results analysis, the relative z-score distance measure is used to calculate the similarity between each pair of actors. The principle used is that those closest in four-dimensional space tend to be chosen as the ally. Closeness is derived from the normal application of Pythagoras' theorem applied to relative z-score music space. Figure 3 shows a cumulative frequency graph showing the proportion of position choice against distance. Those closest (z-score distance of less than 0.8) are 100% likely to be chosen but at distances approaching a relative z-score of 4 all ally positions are accounted for. Thus at this high inclusive distance all people that can be selected are included.

In Figure 3 the blue area occupying the bottom left quadrant of the graphic shows the likelihood of a first ally being selected as the inclusive distance is extended. When the evaluative distance is at its closest, a z-score of 0.4, then the likelihood of 1st choice is a certainty. As the distance increases the likelihood begins to even out between the 1st, 2nd (maroon) and 3rd (yellow) choices because more people are being included. The position of no correlation between closeness and ally choice would give the result that all ally choice positions would be equally likely for all relative z-scores.

5. Model Running Results Analysis

Seven experimental runs of our model were done utilising the start conditions of the seven actual experiments with people.

A model was developed to simulate how one participant can acquire another's internal view of music through conversation and inference using simplified questions without directly asking about their scale. We set the model up by using for each subject the actual initial scores expressed during our experiments. We refer here to the model of a subject as an 'actor' since the software is playing the part of an experimental subject in this scenario. We organise the model so that the actors are in the same groupings as the human subjects.

In the model the actors enter into a conversation with the other actors in the group with a limited questioning repertoire viz. only being able to ask if they like a particular piece more than another piece, and which they liked least or most.[2] The choice of to whom the question is directed, what music is compared and what question is framed, is left to the model of the actor's choice. The choice is based upon the degree of uncertainty of another actor's view and the purpose of the conversation. This purpose is to reduce the actor's degree of uncertainty about its view of all the other actors in the group. The mechanism uses game theory as described by Addis and Gooding (1999, 2004). In the model the other actors don't have access to individual conversations, i.e. they are not listening in; they only know about the answers to their own questions.

Each actor has their own separate sub-model for each of the other actor's internal views and these sub-models are modified according to the answers they receive in response to their questions. Based on these four sub-models (one is the actor's own scale and the other three are scales from its belief model of the other actors), we can calculate the distance measurement for each actor as perceived by any other actor through their sub-model. In comparing the actual mentor choice order with the

[2] Ordinal relationships are one of Wittgenstein's objects as defined in his *Tractatus*. This would suggest that denotational semantics is being used in this case despite the lack of direct observation. We will be discussing this elsewhere.

model distance, we should be able to get similar results to those above (Figure 3) if the actor's belief model does represent other's internal view to some degree.

Model Data Input and Results Analysis

Seven sets of actual experimental data are used as input. For each set, the model is run 300 times to get the actors' expected belief model result. The following figure (Figure 4) shows how Actor 1's view of Actor 2 changes through the conversation.

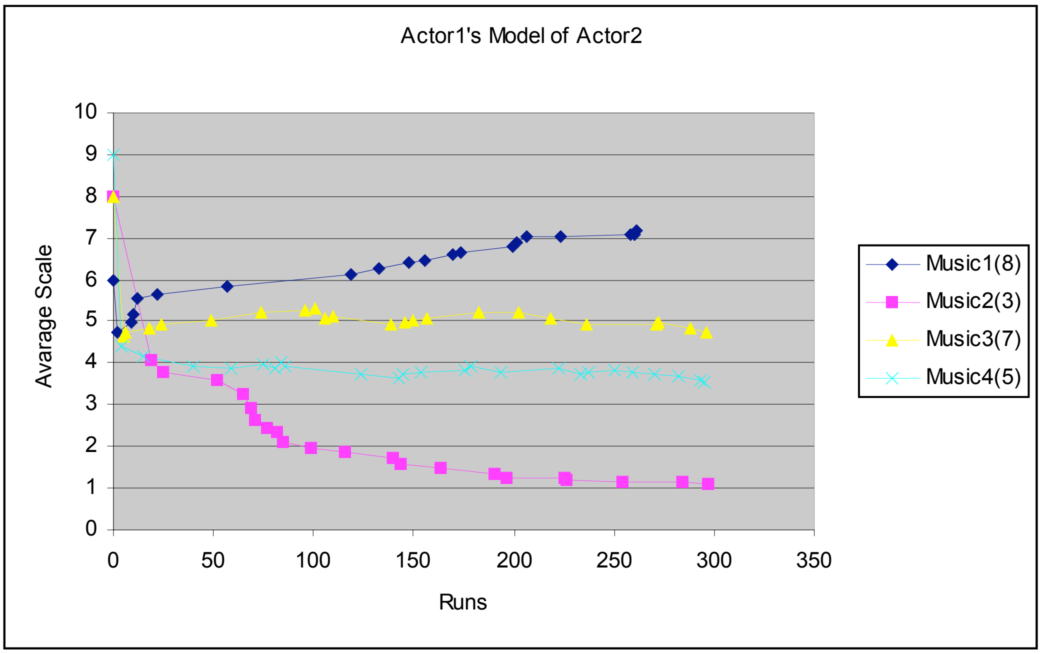

Figure 4: How Actor 1's model of Actor 2's evaluative scale changes over 300 runs

It can be seen that as the model runs and actor 1 asks questions of actor 2, actor1 homes in on actor 2's order of evaluation but not exactly how much it 'likes' each piece. The realisation of sequence occurs fairly fast, by less than 30 runs, and thereafter actor 1's belief only strengthens as to sequence and broadens as to the absolute scaling he believes to be the view of actor 2. (See Figure 4 where bracketed numbers in the key are the actual values for actor 2). Note that actor 1 does not always ask actor 2 at each cycle and in this case only asks actor 2 about ten question before deciding its order. The run of questions (see Figure 4, Music 2 between 50 and 100 runs) is caused, in part, by the relative uncertainty actor 1 has about actor 2 compared with the other actors in the group.

Actor 1's belief model of actor 2 is given in table 1 below. This includes 4 ranges of belief (series 1 to 4) for each of the four music objects. The belief concerns the scales 0 to 9 such that actor 1 has some expectation as to each scale position concerning a piece of music for each actor. The sum of these beliefs for a series adds up to one since the actor must place the music somewhere on the scale. There are three important measures derived from the ranges of belief:

$$\text{Indifference, } I(x) = \text{Anti-Log}_2 \left(- \textstyle\sum_x p_x.\log_2 (p_x)\right)$$

Where p is the belief of x for a given actor's perception of another actor's view of a piece of music. I(x) represents the value that a belief would need to have if, for the same level of overall uncertainty (entropy), the belief value were to be equal for all hypotheses (scale positions). Under this hypothetical situation all the hypotheses (scale values in this case) would be indifferent to each other. We take this level of indifference to be a threshold above which the hypotheses are considered 'believed' and below which they are considered 'disbelieved'. This threshold is dynamic and tends to become higher as more hypotheses' belief values approach zero.

$$\text{Expectation } E(x) = \sum_x p_x.x$$

Expectation E(x) is the expected or average value of the scale over an imaginary time period (n), which in this model is taken to be in the order of four events. It is 'imaginary' because it is in practice calculated in terms of an accumulation of effects such that events occurring further back in time have an exponentially decreasing weight on the current value. The equivalent time period in this case is in the order of four events. The consequence of making this time window larger is to reduce the response of the belief value to change, so we have:

$$\text{Flexibility } f(p_x) = 1/n$$

Table 1: Actor 1's range of beliefs of actor 2 after 300 runs

x / music	0	1	2	3	4	5	6	7	8	9	Indifference I(x)	Expected E(x)
Series 1	0.01	0.01	0.02	0.03	0.04	0.07	0.1	0.15	0.23	0.3	0.16	7.18
Series 2	0.5	0.24	0.12	0.06	0.03	0.02	0.01	0.01	0.01	0	0.24	1.09
Series 3	0.04	0.07	0.1	0.12	0.14	0.14	0.13	0.12	0.09	0.1	0.11	4.72
Series 4	0.09	0.13	0.15	0.15	0.14	0.12	0.09	0.06	0.04	0	0.11	3.52

Figure 5 shows the differences in belief over the range of scales after 300 model cycles. Considering that the beliefs are independent the smoothness of the curve is comforting. From Table 1 we can see that for Series 1 and 2 the believed values are well defined being 8 or 9 and 0 or 1. For series 3 and 4 the believed values are more spread out covering a range of five mid range values in both cases. What is plotted in Figure 4 is the expected value rather than the 'believed' values.

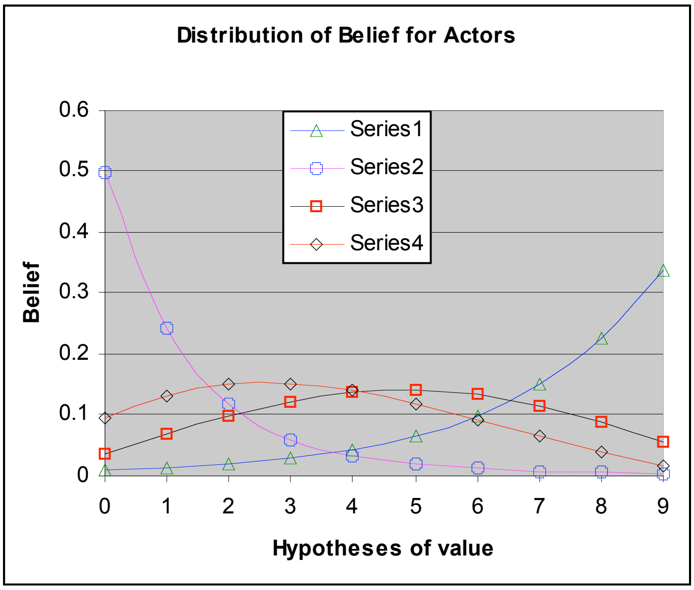

Figure 5. The change of belief of a scale value for actor 2 as perceived by actor 1 after 300 cycles

Table 2: Actor 1's range of expected values for other actors in the group

Series 1	Series 2	Series 3	Series 4	Actor 1 Model Description
6.00	8.00	8.00	9.00	models himself
7.18	1.09	4.72	3.52	models Actor 2
1.61	4.15	6.02	7.81	models Actor 3
0.86	3.5	5.05	6.96	models Actor 4

Table 2 results are subjected to the z-score normalisation process to make them comparable with observed results derived from the experiment. We get the following Table 3:

Table 3: Actor 1's normalised expected values and normalised distances

Series 1	Series 2	Series 3	Series 4	Distance in 4D Space	Actor 1 Model Description
-1.61	0.23	0.23	1.15	0.00	models himself
1.39	-1.38	0.27	-0.28	3.69	models Actor 2
-1.43	-0.33	0.49	1.27	0.65	models Actor 3
-1.45	-0.27	0.43	1.29	0.57	models Actor 4

Then the evaluation distances are calculated by plotting these data sets in four-dimensional space. The distance between each pair of points in 4D space is given indicating how closely allied the other actors are placed in terms of music order in the model. It should be noted that because each actor has a sub-model of the other actors that are likely to be different then actor 1's perception of distance from actor 2 is likely to be different from actor 2's distance from actor 1. So in comparing the distances derived from the model with the results obtained from the experiments (shown in Table 4) we can make some predictions from the relative z-score distance between two subjects' choice of ally. Note that the results in Table 4 are 500 runs instead of 300 as shown previously. This accounts for the first line of Table 4 being slightly different from the 5th column of Table 3.

Table 4: Results of 500 runs of the model for all the experiments showing relative distance z-scores: the right hand 4 columns are the sequences from the real experimental runs involving people.

Sequence Similarity	**EX.1**	A1	A2	A3	A4	A1	A2	A3	A4
1.0	A1	0	3.71	**0.66**	0.72	0	3	**1**	2
0.0	A2	**3.65**	0	3.33	3.31	**1**	0	2	3
1.0	A3	0.62	3.44	0	**0.56**	2	3	0	**1**
1.0	A4	1.18	3.26	**0.61**	0	2	3	**1**	0
	EX.3	A1	A2	A3	A4	A1	A2	A3	A4
0.0	A1	0	1.67	**2.16**	3.12	0	0	**1**	0
0.0	A2	**1.70**	0	3.28	3.60	**1**	0	3	2
1.0	A3	**2.36**	3.34	0	3.05	**1**	3	0	2
0.0	A4	2.90	3.12	**3.04**	0	3	2	**1**	0
	EX.4	A1	A2	A3	A4	A1	A2	A3	A4
0.5	A1	0	**0.756**	3.44	1.88	0	**1**	0	0
0.5	A2	**0.52**	0	3.34	1.84	**1**	0	0	0
0.0	A3	3.47	**3.32**	0	2.36	0	**1**	0	0
0.0	A4	**1.73**	1.72	2.27	0	**1**	3	2	0
	EX.5	A1	A2	A3	A4	A1	A2	A3	A4
1.0	A1	0	3.63	1.99	**1.61**	0	3	2	**1**
0.5	A2	3.66	0	**2.56**	3.61	0	0	**1**	0
0.0	A3	2.21	**2.62**	0	3.03	2	**1**	0	3
1.0	A4	**1.29**	3.58	3.01	0	**1**	3	2	0
	EX.6	A1	A2	A3	A4	A1	A2	A3	A4
0.0	A1	0	**3.85**	3.21	2.16	0	**1**	3	2
1.0	A2	3.84	0	**3.01**	3.37	3	0	**1**	2
0.0	A3	3.38	**2.85**	0	3.46	3	**1**	0	2
0.0	A4	2.11	3.17	**3.44**	0	2	3	**1**	0
	EX.7	A1	A2	A3	A4	A1	A2	A3	A4
0.0	A1	0	2.17	3.41	**3.30**	0	2	3	**1**
0.0	A2	1.84	0	**2.11**	3.97	2	0	**1**	3
1.0	A3	3.44	**2.02**	0	3.10	3	**1**	0	2
0.0	A4	3.33	3.97	**3.63**	0	2	3	**1**	0
9.5/24 = 0.40									

6. Assessing the results

We can measure the predictive power of the model in terms of the improved information over the null hypothesis. The null hypothesis is where the choice of order is made randomly compared with model differences. The model has simulated a conversational process that allows internal sub-models to be constructed that provides an actor centred view. In Table 4 we have scored a successful prediction of order by assigning 1.0 and a partial order as 0.5 only where, due to lack of information, there is 0.5 probability of the answer being correct (rounded to 2 decimal places). All other orderings have been scored 0. So we have:

- Probability of guessing order correctly if random = 1/3! = 1/6 =0.17
- Random Hypothesis Entropy = $-\mathrm{Log}_2(0.17) = 2.56$
- Observed Entropy = $-\mathrm{Log}_2(0.40) = 1.32$

The numeric value of the prediction from the model is the difference of the two hypotheses. This is about 1.2 bits, which means that you roughly double your chance of guessing correctly by running and using the model.

Using the cumulative binomial probability calculations (Feller 1968) to assess the confidence of these results we find that the probability of getting eight or more correct sequences is 0.035444 or less than 4% of the time. For nine or more correct sequences it drops to 0.01176 or slightly greater than 1% and ten or more it is 0.003339 or a bit higher than 0.3%.

Ally Choice

The results seem to indicate that the frequency of ally choice order relates to the relative z-score distances calculated. The complete sequences were not always given but where a particular position is given we have included this in Figure 6 and Figure 7. We can compare the shape of the cumulative graphs for relative z-score distance for each position (first, second and third) for both the observed and the model. The model is on the left and the observed results on the right:

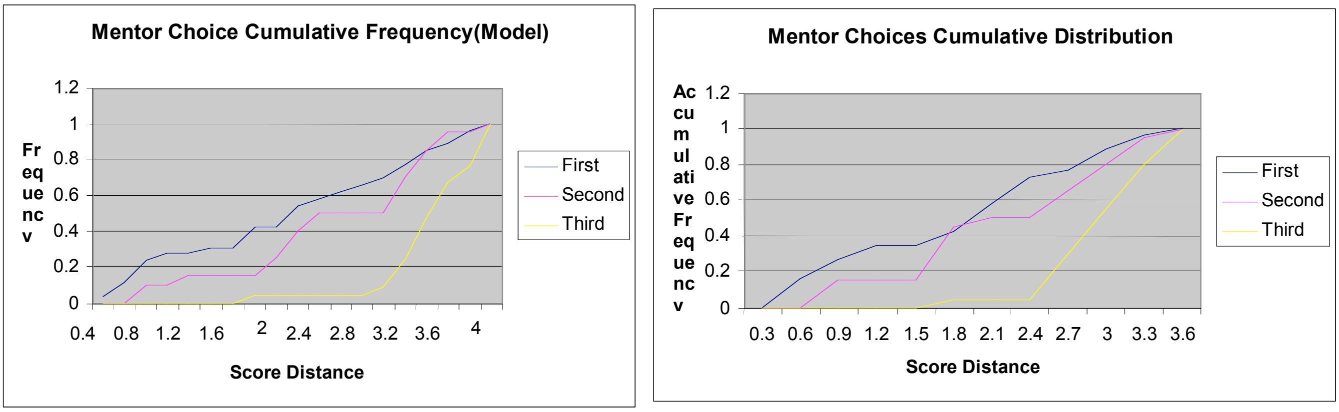

Figure 6: A comparison between observed and model ally choice cumulative probabilities

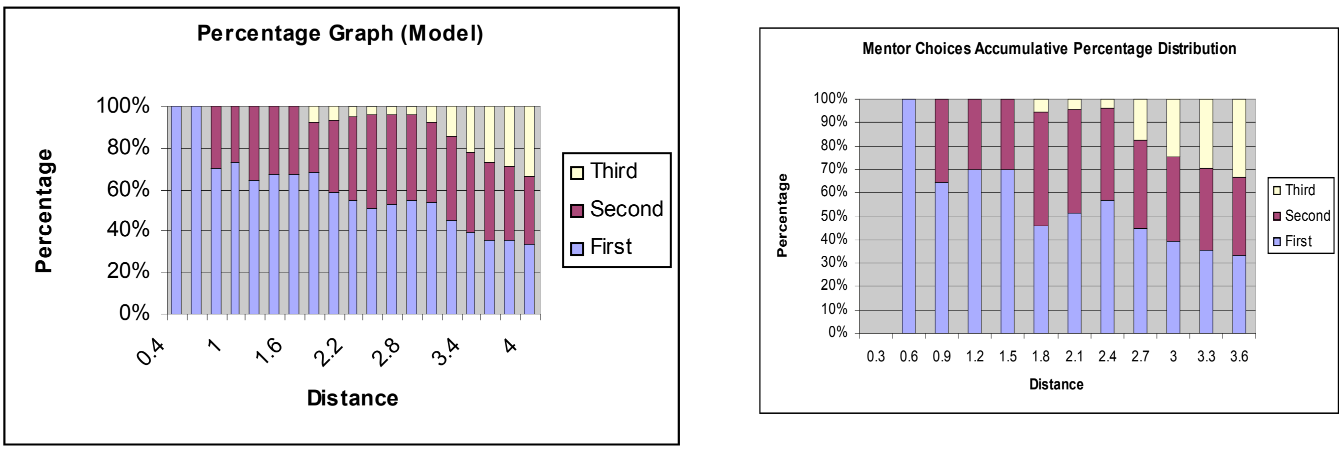

Figure 7: Results of ally choice as structured bar charts: compare Figure 3

Figure 7 shows a similarity between both observed and model results.

7. Conclusion

Our analysis of the experiment with people involving music and our computer model is still to be completed. The experiments have shown that the ranking choice of music seems to follow no pattern and that for each person each piece of music can be treated like a unique dimension. It may be that because all the pieces of music were 'classical' and therefore of the same type that grouping of choice was masked. The most significant information that has been gathered has been the ally choice. It is possible to surmise that such information has a valuable role in the survival, or at least well-being, of individuals since others' experiences become valuable surrogates to one's own experiences. This is a kind of metaphor in that the judgement of a close ally is taken to be representative of one's own experience.

The other result that came from the experiments was that the correlation between group results and individual results was significant. However, there is enough variation in the final decision to suggest that there are factors to be considered other than just an average weight drawn from everyone's opinion. Similarly, the choice of ally shows sufficient variation from the choice of 'like-mind' to suggest that there are reasons for choice other than selecting a perceptual substitute. For clues to this we will need to look at the details of the group discussion.

The model has shown similar behaviour to the experimental observations. Details of model discussions will need to be explored to see what causes clustering of interrogation and whether this is reflected in the observed behaviour of people. Further, we can propose that some of the observed variation is the effect of dominance by individuals. Such dominance can be extracted from our scale data and used to make better predictions of the group decisions. If such dominance is supported then the model could be modified to take this into account and thus give a better account of the group discussions and ally choice.

We are only in the foothills of modelling human behaviour. The advantage of considering music is that the perception of it is purely subjective. There can be no argument or logic as to the nature of the experience for each individual. Only they know what they felt. This neutral stance means that all the mechanisms of language and communication has got to go into communicating these perceptions. The detection of perceptual allies seems valuable for extending ones own experiences and this should lead to using shared metaphors; metaphors that are initially found through direct shared experiences and later from shared perceptions (Lakoff and Johnson 1980, Lakoff 1986). The modelling opens up the possibility of checking the consistency of how this unanchored conversation can drift without loss of communication. In here somewhere, is the suggestion of how ontological changes, e.g. changes in conceptual boundaries, can be made because of the fluidity of the inferential language. In here is a clue to the mechanism of insight and originality. The progress of human thinking and experience seems to be related to our method of communicating between ourselves; our knowledge seems to exist between us more than it exists within each of us in isolation.

8. Acknowledgements

We would particularly like to thanks Dr David Salt for his detailed statistical analysis of the correlations between the musical extract raw scores. Thanks are also due to Prof David Gooding for directing us to the ideas of metaphor and for his work on the contributing belief system. Finally our thanks and appreciation to Jan Townsend-Addis for both her programming and mathematical support.

9. References

Addis, T. et al (2004) *The Abductive Loop: Tracking Irrational Sets* (this publication)

Addis, T. and Gooding, D. (1999), *Learning as collective belief-revision: simulating reasoning about disparate phenomena,* Proceedings of the AISB'99 Symposium on Scientific Creativity, ISBN 1 902 956044.

Addis, T. and Gooding, D. (2004) *Methods for an Abductive System in Science* MBR04 Proceedings (forthcoming)

Billinge, D. (2000) *An Analysis of the Communicability of Musical Predication: a feasibility study for artistic decision support systems*, PhD Thesis University of Portsmouth
Billinge, D. & Addis, T. (2004) *Music to our ears: a required paradigm shift in Computer Science* European Conference on Computing and Philosophy, University of Pavia, Italy
Billinge, D. & Addis, T. (2003) *The Functioning of Tropic Communication: A Mechanism for Consistent Figurative Descriptions of Artistic Effect*. AISB'03 Symposium on AI and Creativity in Arts and Science, University of Wales at Aberystwyth
Feller, W. (1968) *An Introduction to Probability Theory and its Applications*, Third Edition Volume 1. J Wiley & Sons, NY
Gooding, D. C. and Addis, T. R. [in press] *Modeling Scientific Experiments as Mediating Models,* - given at MBR04
Hewitt, C (1979) *Control Structures as Patterns of Passing Messages*, in Artificial Intelligence: An MIT Perspective. Vol 2, Published by MIT Press.
Kuhn, T. S. (1985) The Essential Tension: Selected Studies in Scientific Tradition and Change, University of Chicago Press
Lakoff, G. and Johnson M. (1980). *Metaphors We Live By*. University of Chicago Press.
Lakoff, G. (1986). *Women, Fire, and Dangerous Things*. University of Chicago Press.
Popper, K. (1959) *The Logic of Scientific Discovery* 10th Impression 1980, Hutchinson.
Stepney, S., Braunstein, S., Clark, J., Tyrrel, A., Adamatzky, A., Smith, R., Addis, T., Johnson, C., Timmis, J., Welch, P., Milner, R., Partridge, D. (2004) Journey: Non-Classical Philosophy - Socially Sensitive Computing in Journeys Non-Classical Computation: A Grand Challenge for Computing Research, 18 May 2004, http://www.cs.york.ac.uk/nature/gc7/newcastle.htm
Weiner, P.P. (1966). *Charles S. Peirce: Selected Writings*. New York: Dover.