# Modeling Coordination Problems in a Music Ensemble: Some Logical and Game Theoretic Considerations

Søren R. Frimodt-Møller[1]

[1] University of Southern Denmark, Institute for Philosophy, Education and the Study of Religions, Philosophy, Campusvej 55, DK-5230 Odense M, Denmark
srfm@ifpr.sdu.dk
http://www.orkesterfilosofi.dk

**Abstract.** This paper considers in general terms, how musicians are able to coordinate through rational choices in a situation of (temporary) doubt in an ensemble performance. A fictitious example involving a 5-bar development in an unknown piece of music is analyzed in terms of epistemic logic, more specifically a multi-agent system, where it is shown that perfect coordination can only be certain to take place if the musicians have common knowledge of certain rules of the composition. We subsequently argue, however, that the musicians need not agree on the central features of the piece of music in order to coordinate. Such coordination can be described in terms of Michael Bacharach's theory of variable frames as an aid to solve game theoretic coordination problems.

## 1 Introduction

Consider the following situation in a music ensemble: We have three players, for the sake of desirable connotations let us denote them "the oboe", "the violin" and "the cello". They are playing a new piece of scored music that is hence not part of their individual heritage as musicians.[1] Let us for simplicity consider 5 bars in this score, denoted bars 1-5 (although they may be thought to occur at a later occasion than the beginning of the piece). Still for simplicity, we decide that in these bars the three players each have two possible actions. An action is in this context a phrase to be played within a bar. To echo the theory of multi-agent systems as presented by Fagin et al

[1] In previous conference presentations ("How Do Musicians Reach an Agreement? The Ensemble as a Multi-Agent System" at *Workshop on Deontic Logic*, Roskilde University, November 9, 2007 and participation in "Workshop on Academic Writing" at the annual graduate conference arranged by the Danish Research School in Philosophy, History of Ideas and History of Science, Sandbjerg Estate, December 7, 2007) I have described how the same sort of doubt may arise in a known piece of music, for instance bars 4 to 8 in Schubert's *Unfinished Symphony*. I have, however, found that my audience is less likely to accept that it can be problematic for skilled musicians to coordinate in such a (presumably) familiar context, therefore I have generalized the example.

(1995)[2] let us call the set of possible actions for a player i the set of possible *local states* for that player, $L_i$. We now define:

$L_{oboe}$={phrase1,phrase2}
$L_{violin}$={phrase3, phrase4}
$L_{cello}$={phrase 5, phrase6}

(To make the example more in accordance with reality, we could add a state Λ to each of the sets $L_i$, denoting that the player does not play anything. We will, however, not consider cases where such behavior is involved here, and therefore we omit these possible states. We could also have decided on a more general definition of a state to include any sort of event and subsequently added a set $L_e$ of possible states for the environment, where we could have placed events external to the ensemble that may affect there actions, such as "a truck passes the concert hall". But due to the fairly short length of this paper we only consider the behavior of our three players in their interrelations.)

Now, according to the score, the three players are supposed to play their phrases in a rather staircase-like development: In bars 1-2, the oboe is supposed to play phrase 1, the violin phrase 3 and the cello phrase 5. In bar 3, the oboe is supposed to play phrase 2, the violin phrase 3 and the cello phrase 5. In bars 4-5 the oboe returns to playing phrase 1, but the violin plays phrase 4 in bar 4 and then returns to phrase 3 in bar 5, whereas the cello continues playing phrase 5 in bar 4 and then plays phrase 6 in bar 5. The situation is illustrated in Table 1.

| | Bar 1 | Bar 2 | Bar 3 | Bar 4 | Bar 5 |
|---|---|---|---|---|---|
| oboe | Phrase 1 | phrase 1 | phrase 2 | phrase 1 | phrase 1 |
| violin | Phrase 3 | phrase 3 | phrase 3 | phrase 4 | phrase 3 |
| cello | Phrase 5 | phrase 5 | phrase 5 | phrase 5 | phrase 6 |

*Table 1.*

Intuitively, the violin should wait for the oboe to play phrase 2 and then play phrase 4 at the following bar. The cello should wait for the violin to play phrase 4 and then play phrase 6 at the following bar. Now consider what happens, if the oboe plays phrase 1 three times in a row. The violin might either think "too bad for her, I'm proceeding to bar four anyway, or else the cello will not know what to do" or "I'd better wait for the oboe to play her phrase 2 and then interpret that bar as bar 3 and the following as my bar 4." But what will he think? This depends on how important he thinks the development in the phrasing of the oboe is in comparison with the development of his own phrasing not to mention that of the cello. Let us say that the violin chooses to pursue the second tactic, namely resume playing his phrase 3 until he hears the oboe playing phrase 2. What will the cello think? The cello might think, "The oboe and the violin have both got it wrong, but that is not my problem, I am going for the fifth bar in this development with my phrase 6 as planned, then they can adjust to what I am doing in the following bar." But she might also think "Oh, we should probably wait for the oboe to commence her phrase 2 and then continue the development as if that bar was

---

[2] In which the theory is used to describe problems involving both communication and coordination such as the Problem of Coordinated Attack (see i.e. 109-122 and 190-199)

bar 3." (She might actually also think "Never mind the oboe, I will wait for the violin to play his phrase 4 and then play phrase 6 at the next bar", but this will amount to the same line of action although the intention is different.) What she thinks depends on whether she thinks her own voice or that of the oboe (or, for completeness, that of the violin) is the most important in this section of the piece. Our troubles do not end here. The oboe might also be considering what to do next, e.g. wonder whether she should just think "Oh no, I blew it, but too bad, I just have to continue according to the score" or "the other musicians are waiting for my phrase 2, so I should play phrase 2 to get things going." As with the other two, what she chooses to do depends on how she conceives of the composition.

The score might quite probably give some clear normative guidelines as to what is the most important in the composition. (We think of the composition as something of which the score is an arrangement – in this way we are able to distinguish between different arrangements of a composition, although these arrangements will necessarily deviate in certain aspects from the original score (if any) of the composer.) But it might also be that the question of what the central parts of the composition are is a matter of interpretation, that is, relative to respective musicians.

In the following we will try to elucidate how the three musicians can navigate out of the situation through two different sorts of analysis. The first is in terms of a *multi-agent system* within epistemic logic and considers the case, where a set of guidelines being *common knowledge* in the ensemble will enable the musicians to solve their coordination problem. The second is in terms of Michael Bacharach's idea of the role of *framing* (Bacharach 2006) in game theoretic problems and will address how musicians may be able to coordinate even though they do not have exactly the same opinion of the salient features of the composition.

## 2 The Coordination Problem Analyzed in Terms of a Multi-Agent System

Let us pick up our definition above of the local states of the players. Fagin et al (1995, 110-111) think of a state as an *information state*, that is, a state that contains information. Strictly speaking, if we want to follow this line of reasoning, we should add a number of possible local states for a player i containing information not only about what i is playing now, but also about what the other players are playing, and what everyone was playing at previous bars. In our example here, however, we assume that all players actually hear everything that happens, and that they have perfect memory. We therefore assume that everyone is always aware of what is happening at all local states, and for the sake of simplicity we choose to model the information state of player i as only containing information about the action of i at a given time.

We now define a global state, $G = (s_{oboe}, s_{violin}, s_{cello})$, where $s_i$ is the state for the player i (in our example, the phrase that the player is playing). Intuitively $G$ expresses some situation where each of the players is playing a specific phrase from his or her respective set of possible states. We thus have a set of possible global states, $\mathcal{G}^{\text{ensemble}} = L_{\text{oboe}} \times L_{\text{violin}} \times L_{\text{cello}}$ (the Cartesian product of all the sets of possible local states).[3]

We would like to model *G* as a function of time. For the present purposes we think of time as being discreet and introduce a point in time m, $m \in \{0,1\ldots.\}$. This is quite convenient because it allows us to think of steps in time as being synchronous with and equal to the length of developments from one bar to another, which is exactly what we will do. We define *a run* to be a description of how the global state develops through time, more precisely, the global state as a function of m: $r(m) = (s_{\text{oboe}}, s_{\text{violin}}, s_{\text{cello}})$, such that r(0) is the initial global state, r(1) the next global state etc. We now define a multi-agent system $\mathcal{R}^{\text{ensemble}}$ over $\mathcal{G}^{\text{ensemble}}$ as a set of runs over $\mathcal{G}^{\text{ensemble}}$. A point (r,m) is the time-point m in the run r. We say that (r,m) is a point in the system $\mathcal{R}^{\text{ensemble}}$, if $r \in \mathcal{R}^{\text{ensemble}}$. $r_i(m) = s_i$, so that $r_i(m)$ is player i's local state at the point (r,m).

Before we can analyze our coordination problem above, we need to define what it means for a player to distinguish (or not be able to distinguish) between two global states:

- Let $s = (s_{\text{oboe}}, s_{\text{violin}}, s_{\text{cello}})$ and $s' = (s'_{\text{oboe}}, s'_{\text{violin}}, s_{\text{cello}})$ be two global states in $\mathcal{R}^{\text{ensemble}}$. We say that player i cannot distinguish s from s´, notated $s \sim_i s'$, if player i has the same state in s and s´, in other words if $s_i = s'_i$ (An important thing to notice here is that we think of a state not only as the action of a player but as a situation in which the player chooses the action that has given rise to the label of that state.)
- In accordance with this we say that player i cannot distinguish between two points (r,m) and (r´,m), $(r,m) \sim_i (r',m)$ if $r(m) \sim_i r'(m)$, in other words if $r_i(m) = r'_i(m)$

In the epistemic logic of multi-agent systems described by Fagin et al (1995), the notion of indistinguishability is used to define the operator $K_i$, which in our case would intuitively mean "player i knows that…" In this example we will not need to make statements about the players' knowledge of propositional facts, only their awareness of the global state and its relation to other global states, hence we omit the definition of the $K_i$-operator.[4]

With these formalities in place we can now describe the stepwise development of our coordination problem formulated as the system $\mathcal{R}^{\text{ensemble}}$. As hinted at before, we take the time variable m to be a stepwise development of one bar length. In a case where all three musicians follow the score perfectly (a specific run in $\mathcal{R}^{\text{ensemble}}$ which we choose to label $r^{\text{score}}$), m should therefore be perfectly synchronized with the bar numbers such that the global states develop in this way:

---

[3] These formalizations are identical with the definitions given by Fagin et al (1995, 111-121). The following formalizations are my versions of definitions given in the same pages, only adapted to my own example.

[4] We also omit describing the Kripke-structure associated with the interpreted system $\mathcal{I}^{\text{ensemble}}$ (the system $\mathcal{R}^{\text{ensemble}}$ along with an assignment of truth values to all propositions that occur in the system for each state in the system) as we will not have need for it here. For a discussion of this aspect of the semantics of multi-agent systems, see Fagin et al (1995, 117-118).

$r^{score}(1)$ = (phrase 1, phrase 3, phrase 5)
$r^{score}$ (2) = (phrase 1, phrase 3, phrase 5)
$r^{score}$ (3) = (phrase 2, phrase 3, phrase 5)
$r^{score}$ (4) = (phrase 1, phrase 4, phrase 5)
$r^{score}$ (5) = (phrase 1, phrase 3, phrase 6)

Now let us look at a case where the oboe forgets to play phrase 2 at bar 3. A number of different runs might then occur in which the first three steps would be

$r^{late(u)}(1)$ = (phrase 1, phrase 3, phrase 5)
$r^{late(u)}(2)$ = (phrase 1, phrase 3, phrase 5)
$r^{late(u)}(3)$ = (phrase 1, phrase 3, phrase 5)

(u should be read as a variable that can be substituted for a specific label.)

If phrase 2 and phrase 4 are strongly dissonant, the musicians would probably want to avoid a scenario where the two phrases occur at the same bar. In other words we would e.g. like to avoid the run $r^{lateoboe}$ where

$r^{lateoboe}(4)$ = (phrase 2, phrase 4, phrase 5)

Suppose that the violin chooses to wait for the oboe instead of proceeding according to the score. Then we would have a run $r^{lateviolin}$ where

$r^{lateviolin}(4)$ = (phrase 2, phrase 3, phrase 5)

But this run might continue in two different ways: One in which the cello adjusts to the other players and do not play phrase 6 until the bar after the violin has played phrase 4 (which would be a bar beyond our current example), and another in which the cello proceeds according to the score, that is, where we end up with

$r^{lateviolin}(5)$ = (phrase 1, phrase 4, phrase 6).

Of course for all 1,307,674,368,000 possible deviations from the score, there is the possibility that everyone, including the player(s) with erroneous phrases, try to keep following the score as closely as possible by playing the "right" phrase according to the score at the next m (thus interpreted as a bar number). For simplicity, we will not try to describe this general case formally here. For convenience, we may, however, add a run describing the situation where the oboe forgets to play phrase 2 at m=3, but where everyone, including the oboe, continues according to the score:

$r^{latescorevar1}(4)$ = (phrase 1, phrase 4, phrase 5)
$r^{latescorevar1}(5)$ = (phrase 1, phrase 3, phrase 6)

And we can add a run describing the situation where the violin considers his own phrase 4 more important than the oboe's phrase 2, where he nevertheless forgets to play this at m=4, but where everyone continues according to the score at m=5:

$r^{latescorevar2}(4)$ = (phrase 1, phrase 3, phrase 5)
$r^{latescorevar2}(5)$ = (phrase 1, phrase 3, phrase 6)

To nearly complete the picture,[5] let us describe the case where the oboe forgets to play phrase 2 at m=3, but plays phrase 2 at a m=t, t>3, where the violin chooses to wait for the oboe and reinterpret the bar where the oboe plays phrase 2 as bar 3 according to the score, and where the cello likewise interprets the bar where the violin plays phrase 4 as bar 4 according to the score:

$r^{lateviolinwaits}(t)$ = (phrase 2, phrase 3, phrase 5)
$r^{lateviolinwaits}(t+1)$ = (phrase 1, phrase 4, phrase 5)
$r^{lateviolinwaits}(t+2)$ = (phrase 1, phrase 3, phrase 6)

(We could also describe a situation where the violin does not wait for the oboe, but where the cello will wait for the violin. This is, however, not of relevance to our analysis of the example at this point.)

Now we can identify and formalize the situations of doubt the three players may experience when the oboe forgets to play phrase 2 at bar 3. At m=3, the violin does presumably realize that the other players are no longer proceeding according to $r^{score}$, but he does not know (in our current description of the full situation) whether the other players are proceeding according to $r^{lateoboe}$, $r^{lateviolin}$, $r^{latescorevar1}$, $r^{latescorevar2}$ or $r^{lateviolinwaits}$. Formally $(r^{lateoboe}, 3) \sim_{violin} (r^{lateviolin}, 3)$; $(r^{lateoboe}, 3) \sim_{violin} (r^{lateviolinwaits}, 3)$, $(r^{lateoboe}, 3) \sim_{violin} (r^{latescorevar1}, 3)$, $(r^{lateoboe}, 3) \sim_{violin} (r^{latescorevar2}, 3)$ and $(r^{lateoboe}, 3) \sim_{violin} (r^{lateviolinwaits}, 3)$. So how can he ever know what would be the appropriate way to proceed at m=4, except by picking a choice at random? In fact, this situation is the case for all of the players, hence $(r^{lateoboe}, 3) \sim_i (r^{lateviolin}, 3)$; $(r^{lateoboe}, 3) \sim_i (r^{lateviolinwaits}, 3)$, $(r^{lateoboe}, 3) \sim_i (r^{latescorevar1}, 3)$, $(r^{lateoboe}, 3) \sim_i (r^{latescorevar2}, 3)$ and $(r^{lateoboe}, 3) \sim_i (r^{lateviolinwaits}, 3)$, because everyone has the same (local) state at m=3 no matter which of the runs is executed. (Strictly speaking, it is rather unlikely that any of the players should consciously choose to follow $r^{lateoboe}$ or $r^{lateviolin}$, but we will return to the discussion of what strategy a player is likely to choose later in this paper.)

In order for the players to be able to make a rational choice of what to play at m=4 and onwards, they must either have common knowledge of some rule that clearly states which of the runs is being executed, or they must have some way of getting about the problem of disagreement on the character of the run. The latter set of options is explored in the sections below. The aforementioned rule could be stated as a rather strict obligation to wait for the oboe's phrase 2 and then proceed according to $r^{lateviolinwaits}$, but a formalization of this will necessitate an introduction to deontic logic

[5] Because the players can all hear each other, we have reason to believe that the oboe will understand that the other players have proceeded past bars 4 and 5 in the score, once she hears them play phrase 4 and phrase 6 respectively in succession.

as well as temporal operators, for which we do not have the sufficient amount of space here. We will, however, dwell for a moment on the topic of what it means for such a rule to be common knowledge among the players.

## 2.1 Common Knowledge of Rules and Its Implications for the Ensemble

A statement p being common knowledge in a group G, notated $C_Gp$, entails informally that everyone in the group knows p, and that the entire group is somehow aware of p's being known by everyone and being expected to be known by everyone. The formal representation of $C_G$ in terms of the operator $E_G$, meaning "everyone in G knows that" is debated[6], but all theories grant that $E_Gp$ can be deduced from $C_Gp$, and hence that $K_ip$ (i knows that p) can be deduced from $C_Gp$, for all $i \in G$.

Intuitively it should not be surprising that common knowledge in the group of a rule is required in a situation where coordination depends on the group members following the rule. In our example above, it is not enough that everyone in the ensemble knows that a rule p holds, if someone is in doubt whether the other ensemble members know that rule p holds. (We are of course still assuming that the players have no way of communicating that they follow p during a performance.) On the other hand, once p is common knowledge in the ensemble, that is, once it is part of the collective consciousness of the ensemble, it is safe to entail that everyone in the ensemble knows p. And since p is a rule that states what the ensemble should do when deviating from the score, knowing this rule combined with knowing that everyone else knows it and assumes that everyone else knows it, results in the individual ensemble member following the rule, thus ensuring coordination.

The idea of the ensemble being collectively conscious of a coordinating rule p, however, amounts to an idea of the ensemble having the same opinion of the salient features of the composition. Remember that in our description of the piece of music, we do not know whether the oboe, the violin or the cello has the most important role in the passage. It might be that the voice of the oboe is not only the initiator of a stepwise development in the voices but also an indispensable part of this development, for example if the sequence phrase 2 – phrase 4 – phrase 6 constitutes a melody that simply for the purpose of a fun effect has been distributed onto three different voices. But it might also be that the oboe's phrase 2 is just like a small prologue to a theme that actually begins with phrase 4 in the violin, and that phrase 4 for some reason is tightly knit to a rhythmic structure that develops over bars 1-3. A similar situation could be the case for the cello, if the oboe and the violin are merely adding small fills to the last of four bars that naturally precede phrase 6 in the cello. In any of the three cases, if we could point to a rule that, if common knowledge in the ensemble, would ensure safe conduct in the situation of doubt, this rule would indirectly be a statement of the compositional features to be regarded as salient by every musician. In other words, this account of coordination in the ensemble leaves no possibility of disagreement with

[6] Fagin et al (1995) discuss at least two different interpretations of the notion, one in terms of a possible infinite iteration of the $E_G$-operator (23-25), another in terms of sets of information states in so-called Aumann structures (38-41).

respect to the interpretation of the composition. For a programmer simulating an ensemble as one virtual accompanist to one live soloist, this is not a big issue. We would generally like an accompanist that, at the worst, is only in disagreement with the soloist, not with itself also. For someone modeling the interactions of several independent players, modeling players with different initial perspectives on the music is, however, very important.[7]

In the following sections, we will examine what can be done for a formal description of ensemble coordination without imposing a structure where everyone has to have the same idea of the salient features of the composition.

## 3 Game Theory with Variable Frames

A great deal of effort has been put into explaining how people are able to coordinate in games where two or more players (here understood as players of the game, not musicians) only receive a payoff, if they are able to simultaneously choose the same of a number of options. For instance, in the introduction by Natalie Gold and Robert Sugden to Bacharach (2006), we find the example of "Three Cubes and a Pyramid" (19). In this game two players have to choose the same out of four objects, a red cube, a blue cube, a yellow cube and a green pyramid. From an objective point of view, the probability that the two players coordinate on the same object is just 0.25, because there are 16 possible combinations of actions of the two and 4 possible ways they can choose the same object. But experimental studies show (according to Bacharach et al) that people actually tend to be much better at coordinating than that, and that the players tend to choose the green pyramid. The intuitive answer to this question (and the answer given by Schelling (1960, 64) in relation to similar experiments) is that the choice of the green pyramid is somehow more salient than the other options. But why?

First of all, the two players are not just picking at random without taking into consideration how they perceive the game and its four objects. They describe the game to themselves using predicates, and these predicates belong to what Bacharach calls *families* (Bacharach 2006, 14-16) Formally, we define a set $S$ of objects, a set $P$ of predicates and a function $E$ that assigns a (possibly empty) subset of $S$ to each predicate in $P$, such that if $\varphi$ is a predicate, then $E(\varphi)$ is the set of objects $\varphi$ describes (or the *extension* of $\varphi$).[8] If we call the set of objects in the "Three Cubes and a Pyramid"-game $S_{\text{objects}}=\{x_1, x_2, x_3, x_4\}$, and decide that $x_3$ is the green pyramid, we have for instance $E_{\text{objects}}(\text{cube})=\{x_1, x_2, x_4\}$ and $E_{\text{objects}}(\text{pyramid})=\{x_3\}$. If the extension of a predicate has more than one member, such as "cube" in this case, we call the act of singling out one object to which that predicate applies, "picking". If the extension is a singleton, such as the extension of "pyramid", we call the act of singling out the object to which that predicate applies, "choosing". In other words, the players can "pick a cube" or "choose the pyramid" but not "choose a cube" or "pick a pyramid". The

[7] For one of many examples of the efforts being put into achieving alignment of a virtual accompanist's delimitation of what counts as instances of a given piece of music and the interpretation of the same composition by a soloist, see Fox (2007).

[8] This is my rendition of Bacharach 2006, 10-11 and 14-20.

predicates can be arranged in families, understood as sets of predicates, where, if one comes to mind for the player, the other ones will come to mind as well. Hence we can define a shape family, $F_{shape}$= {cube, pyramid…} and a color family $F_{color}$={blue, red, yellow, green…}. We can also define a "generic family" $F_{thing}$={thing}, where $E_{objects}$(thing)={$x_1$, $x_2$, $x_3$, $x_4$}. We might be able to come up with other families and predicates, but let us stop here for the sake of clarity. Now, for each player, we can define a set of families that might come to mind for that player. We call such a set a *frame*. Such a set is a subset of the *universal frame F*, containing all families that can be taken into consideration in the example (thus the universal frame in "Three Cubes and a Pyramid" is $F_{objects}$={$F_{thing}$, $F_{shape}$, $F_{color}$}). Each player assigns to his opponent (we are assuming a game of two players) a probability $v(F_i)$ that the opponent has a family $F_i$ in his frame – this is also called the *availability* of $F_i$. For instance, a player may think that $v(F_{thing})$=1 for his opponent, $v(F_{color})$=0.6 and $v(F_{shape})$=0.8. So, if the player is right in how he considers the availability of the families for his opponent, the probability that the player will look upon the situation as choosing between shapes rather than "non-descript" objects (Bacharach 2006, 16) is 0.8. Because there are three cubes, the possibility of both players coordinating on the same cube if they both decide on the act-description "pick a cube" is 0.33 (1/3), and, if we take the availability of the shape family for granted, the possibility that they coordinate in general is 0.33*0.8 = 0.26. This is only marginally better than the chances of the players when just picking at random. If both of the players decide to "choose the pyramid", however, they have a 1*0.8 = 0.8 chance of perfect coordination, as there is only one pyramid. If we assume that the payoff for coordination is exactly the same no matter what the players agree to do, it seems that choosing the pyramid is a much better option than any other possible act, as the probability that the players coordinate is higher. (Actually, even if we assume that both players assign an availability of 1 to all families in their opponent's frame, "choose the pyramid" will still be the optimal choice. This is because the options "choose the blue", "choose the red", "choose the yellow" and "choose the green" are discarded due to what Bacharach calls the principle of *symmetry disqualification*.[9] This principle roughly entails that if there are two or more predicates from the same family that have exactly the same size of extension in the game, we have no reason for choosing one over the other, and hence we should disregard the family entirely. Another way of putting it in our case is that absence of a stand-out color choice converts the situation to an arbitrary "picking" between act-descriptions issued from the color family where the chances of coordinating are much smaller.)

We will now try to apply some of these ideas to our coordination problem in the music ensemble.

[9] The analysis of "Three Cubes and a Pyramid" is in essence the same as in Bacharach 2006, 19-22, although I have used a slightly different notation utilizing more transparent subscripts for the different variables.

### 3.1 An Analysis of the Musical Coordination Problem in Terms of Variable Frame Theory

Our coordination problem as described in sections 1-2.1 can be interpreted as a coordination game such as the one we have just examined. The object of the "game" in our ensemble is to choose the same strategy as to which phrases should be played at what time and after which phrases. In our example we have roughly four different strategies: The first is where the musicians try to stick to the score as much as possible and disregard mistakes as unfortunate mishaps. The second is where all three musicians regard the oboe's phrase 2 as essential for the continuous development of the piece and thus wait for the oboe, if the oboe is late. The third is where the musicians regard the violin's phrase 4 as essential and therefore disregard the oboe's eventually being late as a source of confusion but wait for the violin to commence phrase 4 before proceeding according to bar 5 in the score. The fourth is where the musicians regard the cello's phrase 6 as essential, so that even if both the oboe and the violin is late, these players will continue playing their phrases 1 and 3 respectively until the cello commences phrase 6. Unless the cello is even later than both of the other players, the first and fourth strategies amount to the same: follow the score and just move on in case of errors. We can thus simplify our example a bit by eliminating the fourth strategy from our considerations. From the cello's point of view, however, the second and third strategies amount to the same line of action: wait for the violin to play phrase 4, then proceed to bar 5. On the other hand, since it is impossible for the oboe to wait for the violin, the oboe considers the first and third strategies similar with respect to her own line of action: in both cases, she should continue according to the score. So, to sum up, the only player for whom it really matters, if the violin's phrase 4 is most important of phrases 2, 4 and 6, is the violin. If we roughen our distinctions a bit, we could say that the violin really faces a problem of choosing between waiting for the oboe's phrase 2 and not waiting for the oboe's phrase 2. Not waiting for the oboe does not rule out the violin being late himself, if he follows the third strategy described above, but that does not change anything for the other two players. We can therefore describe the coordination problem as a game of coordinating on the same choice of strategy, where the two possible strategies are:

"Wait" (meaning "wait for the oboe's phrase 2 (the oboe plays phrase 2 when ready)") and

"Don't Wait" (meaning "do not wait for the oboe's phrase 2 (continue according to the score if the oboe does not play phrase 2 at bar 3)").

The "objective game" in Bacharach's terms (Bacharach 2006, 14), that is, the game without a representation of the players' frames looks like this: Each of the three players have a possibility of 0.25 of coordinating on the same strategy, whether "Wait" or "Don't Wait" (because there are 8 different combinations of strategies for the three players and 2 possible ways they can choose the same line of action). But the objective game only describes the situation as it would be, if the players picked their strategies at random. It is, however, more likely that they describe the two choices to themselves in terms of their qualities. For example a player could say that "Wait" is a more "melodic" solution with respect to phrasing, or s/he could say that "Don't Wait" "keeps the piece going rhythmically" understood such that this strategy is more in accordance with the overall rhythmical structure of the passage. Let us symbolize "Wait" by $x_1$ and "Don't Wait" by $x_2$. Then we can define a family of predicates $F_{rhythm}$={keeps the piece going rhythmically,…}, where $E$(keeps the piece going rhythmically) = {$x_2$}. We can also define a family $F_{melody}$={melodic,…}, where $E$(melodic) = {$x_1$}. If we once again include the generic family $F_{thing}$={thing} where $E$(thing) = {$x_1$, $x_2$}, we have the universal frame $F$={$F_{thing}$, $F_{rhythm}$, $F_{melody}$} for the coordination game. Now, because of the inclusion of $F_{thing}$, a player that has all three of the mentioned families in his frame can decide on one of these act-descriptions: "pick a thing (something)", "choose the option that keeps the piece going rhythmically" or "choose the melodic". I have deliberately simplified the amount of possible choices and predicates in this example, because our example has the complexity over "Four Cubes and a Pyramid" that there is an extra player. Each player assigns two availabilities for a family, that is, one for each of the other players. Let us say that the violin assigns the possibility $v_{oboe}(F_{melody}) = 0.7$ to the case where $F_{melody}$ comes to mind for the oboe, $v_{oboe}(F_{rhythm}) = 0.3$ to the situation where $F_{rhythm}$ comes to mind for the oboe, $v_{cello}(F_{melody}) = 0.6$ to the situation where $F_{melody}$ comes to mind for the cello and $v_{cello}(F_{rhythm}) = 0.5$ to the case where $F_{rhythm}$ comes to mind for the cello. If the violin is right about his estimates and decides to "choose the option that keeps the piece going rhythmically", he has a 0.3*0.5*1 = 0.15 chance of coordinating with the other musicians on this strategy. If on the other hand he decides to "choose the melodic", he has, provided his estimates are correct, a 0.7*0.6*1 = 0.42 chance of coordinating with them on this. This is still not an overwhelming safety, but if we grant that coordination on a strategy is good no matter the strategy, it seems reasonable for the violin to "choose the melodic" because he considers the probability of coordinating with the other two players higher than by picking at random. But does this ensure coordination in the ensemble? This is the subject of the next section.

### 3.2 What Does the Availability of a Frame Show Us?

There are at least two problems that some readers will notice immediately in the analysis above. The first is that it might be that the violin is wrong in his assignments of availabilities to families in the frames of his co-players. The second is that it might be that the other players have a different view of the availabilities of families in each other's frames, thus making the probability assessment even more complicated. It is important to note in connection with these two complications that what we have described above is how a player can rationally make a choice based on his or her expectations of how the other players may be likely to think. Even if the violin is for instance right in his assumption that $v_{oboe}(F_{rhythm}) = 0.3$, this does not mean that it can never occur that the oboe decides to "choose the option that kept the piece going rhythmically". But if his estimates of the availabilities are generally right, and if coordination, no matter the strategy, is still the objective, the violin will be foolish not to go for the strategy that gives him the highest probability of coordination. So the real trouble here is on what basis a player makes his estimates of the availabilities of families in the frames of his co-players. Intuitively, if an ensemble, such as the trio we are considering here, have been working together for a long time, it seems that it would be strange if the players deviated much from each other in their views of the availability of a family in a given player's frame. On the other hand, an ad hoc ensemble of musicians where no one knows each other, might have fairly the same expectations of the availabilities of different families in each other's frames, namely close to 0.5 for all families. The latter situation is, however, not likely to ensure very good coordination because the possibilities for coordination on a strategy will inevitably come out rather low. But both of the mentioned intuitions point to the relevance of musicians "knowing each other" prior to a performance.[10]

Of course, we can still improve the probabilities of coordination in the ensemble by strengthening the common knowledge or "consciousness" of certain rules inherent in the composition. In the above case, the violin would then probably assign the same availabilities to a family in all frames of his co-players. What we wanted to show in our analysis in terms of Bacharach's variable frame theory was, however, that the players might be able to make non-random decisions making coordination quite possible, even if they do not have common knowledge of the rules of the composition but only some expectations of each other's way of perceiving the situation. Such estimates as the one described in section 3.1 does not ensure coordination, but makes coordination more possible than if everyone chooses at random.

## 4 Conclusion

The above analysis is of course simplified but it points to a way of modeling ensemble relations that might be of relevance for researchers in computer music modeling. The idea is that when modeling two or more ensemble players, we should define their (possibly virtual) characteristics as musicians, that is, their musical background such as their tastes, their previous engagements in other ensembles, their cultural heritage etc.[11]. Some of these traits might be quasi-formalized, for instance a strong dependence to follow the score in a rhythmically strict way or a partiality to the execution of central melodic phrases. Depending on the outcome we want, we can make all of these characteristics known to all players, only some of them or none. We can then model how a rational player will navigate in situations of doubt (or, although this requires a different sort of analysis, situations where the musicians deviate from the score on purpose, such as in an improvisation) by computing his possibilities of coordinating with other players on a strategy given his estimate of what they are likely to choose. We can still include normative features of a composition (understood as something of which the score is merely one of many possible arrangements) in the model[12] if we

---

[10] This in accordance with many of the musicians I know who either will not perform with other people without extensive rehearsal or will only perform with people they are familiar with in advance.

[11] I am thinking along the lines of how many chess engines are built. Here the player can choose his opponent among different profiles mimicking different human backgrounds.

[12] Indeed, following Sharpe (2004, 59-60), preserving characterizations of the composition in terms of normativity is essential for the clearest possible delimitation of the boundaries of a piece of music.

want to and define which musicians know these features, but we do not need all of the players to agree on these features in advance for coordination to take place.

In this paper I have tacitly relied on my own experience as a violinist in several ensemble contexts. To achieve more accurate modeling of coordination processes such as the ones I have tried to describe here, it will of course be necessary to conduct experiments and interviews with several more ensemble musicians.

## References


1. Bacharach, M.: Beyond Individual Choice: Teams and Frames in Game Theory. Edited by Gold, N. and Sugden, R. Princeton University Press (2006).
2. Fagin, R., Halpern, J.Y., Moses, Y. and Vardi, M.V.: Reasoning about Knowledge. MIT Press (1995). Page numbers from paperback edition (2003)
3. Fox, C. and Quinn, J.: How to Be Lost: Principled Priming and Pruning with Particles in Score Following. In: Proceedings of the 2007 International Computer Music Conference, vol. I. The International Computer Music Association (2007), 81-84.
4. Sharpe, R.: Philosophy of Music: An Introduction. Acumen (2004)
5. Schelling, T.: The Strategy of Conflict. Harvard University Press (1960).