Toward a Complete Implementation of SeekWell

Eric Nichols

Final Project: Music T561 (Models of Music Cognition)

May 4, 2004

Prof. Eric Isaacson

Table of Contents

1 Purpose

This paper outlines the next stages in the development of the SeekWell model. A simpler version of the SeekWell model has been described in detail, but the full model is still in its infancy. SeekWell, in its final form, is intended to automatically model melodic expectation in simple tonal melodies.

The full model will operate by internally generating its own restricted Schenkerian analysis (or *embellishment hierarchy)* given only a list of pitches; no other auxiliary data such as key information or meter will be provided. Rather, these will be inferred. A major goal of the model is that it will work in a cognitively plausible fashion respecting the temporal aspect of music listening. Further, we expect that many cognitive problems such as that of key determination, meter determination, and embellishment hierarchy generation are mutually dependant and that a unified model can simulate the cognitive processes involved in all of these using the same mechanisms.

First we present a brief history of the SeekWell model including its roots in a project dealing with number sequences. Next we describe a general model architecture – the "parallel terraced scan" – to be used in the SeekWell model proposed here. To motivate the intent of the model further, the intended functionality of the model for two specific examples is described as a sort of thought experiment. Finally, the specific features of the model are detailed and issues are presented to consider during a forthcoming implementation stage.

2 Background

2.1 Seek-Whence

The SeekWell program has its roots in one of the first projects on modeling cognition at Indiana University's Center for Research on Concepts and Cognition (CRCC). This project, Seek-Whence, investigated the perception and completion of patterns in sequences of numbers (Hofstadter and Fluid Analogies Research Group. 1995). A simple sequence would be:

1 – 2 – 3 – 4 – 5 …

This sequence would be expected to continue moving upwards with 6 – 7 – 8 – etc. A more complicated example is:

1 – 2 – 1 – 1 – 3 – 1 – 1 – 4 – 1 – 1 …

Here, the expected continuation relies on mentally grouping the numbers in the sequence as follows:

(1 – 2 – 1) – (1 – 3 – 1) – (1 – 4 – 1) – (1 …

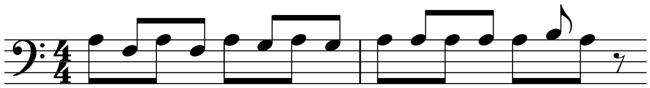
so that the expected continuation is 5 – 1 – 1 – 6 – 1 … A similar but more interesting example is:

2 – 1 – 2 – 2 – 2 – 2 – 2 – 3 –2 – 2 …

Here the long sequence of "2's" is distracting, but when grouped as in the previous example the sequence is easy to comprehend. A final example provides a link to the SeekWell domain:

8 – 6 – 8 – 6 – 8 – 7 – 8 – 7 – 8 – 8 – 8 – 8 – 8 – 9 – 8 …

If we think of the numerals as representing pitches, the sequence extrapolation problem turns into a musical expectation problem like the following:



In this example we hear most of the A notes as a repeating note or pedal while in two cases an A is heard as part of an ascending scalar figure. This gives a hint as to the complexity of both the Seek-Whence and the SeekWell domains: the challenge is to develop a model of how people see or hear patterns like these and then use these patterns to predict or expect what comes next in the sequence. A model for the Seek-Whence domain was investigated by Douglas Hofstadter and Marsha Meredith, leading to a new architecture for modeling the cognitive problems involved (Meredith 1986).

2.2 The FARG Architecture for Fluid Simulation

Because SeekWell grew out of the Seek-Whence domain, we will describe some of the modeling principles used in the Seek-Whence program. Then the SeekWell domain will be discussed in more detail where these principles should prove useful.

2.2.1 Parallel Terraced Scan

A primary characteristic of the Seek-Whence architecture (and indeed, of most of the programs from the CRCC's Fluid Analogies Research Group, or FARG) involves modeling hypothetical subcognitive processes as a collection of many individual cognitive agents. This idea derives from Marvin Minsky's *Society of Mind* (Minsky 1988), which discusses hypothetical mechanisms for thought and portrays the brain not as a monolithic unit but, as the title suggests, as a community of sub-brain components that work together to achieve thought. Another useful metaphor is to think of the brain as an ant colony, where the individual ants correspond with the subcognitive agents.

The "parallel terraced scan" is the modeling algorithm that searches for a solution to a given perception problem by simulating many of these agents working in parallel to solve a problem. For instance, in the Seek-Whence domain, certain agents may act as scouts that examine the number sequence for duplicate numbers. Other agents might look for increasing pairs of numbers. Still others might propose that certain numbers be grouped together, so that yet other agents that deal with groups of numbers could examine groups for similarity.

The benefit of the method is that with many different things going on at once, the search for a "good" way to perceive is more likely to be successful. This "bottom-up" processing results in emergent higher-level behavior that might be argued to resemble human perception. Part of the difficulty with the approach is that the processes being modeled with the collection of agents are assumed to be sub-cognitive, so they are only available to us via introspection indirectly, if at all.

2.2.2 Slippage

Another key concept is that of "slippability" – the natural ability of a concept involved in thinking to be replaced with another related concept if the first concept does not quite fit the situation. For example, in the second Seek-Whence example above, we might have grouped the numbers in this way:

1 – (2 – 1 – 1) – (3 – 1 – 1) – (4 – 1 – 1) ...

However, the lone "1" at the start of the sequence doesn't seem to fit well, which induces some pressure for our representation to change. Indeed, we might notice that we can remove the incongruity by shifting (slipping) the grouping boundaries to arrive at the solution presented earlier:

(1 – 2 – 1) – (1 – 3 – 1) – (1 – 4 – 1) – (1 ...

2.2.3 Temperature

A final component introduced into the architecture is the idea of "temperature", which relates to the "satisfaction" of the program with the current representation. When the temperature is high, the individual agents operate more randomly and the representation is more fluid or breakable. However, as the system settles down on a certain solution, the temperature goes down and it becomes less likely that anything will change. Eventually, the temperature gets low enough that the system stops working, presenting a completed solution.

2.3 SeekWell

SeekWell has been used alternately by Steve Larson to refer to several different but related ideas, and we propose to reuse the term again in this paper in a slightly different manner. SeekWell refers both to a "microdomain" for study, a "game", and several different programs that simulate problem-solving in the SeekWell domain.

2.3.1 The Microdomain

To study fundamental cognitive processes, FARG typically focuses on very restricted domains, or "microdomains". The microdomains lack some of the complexities of real-world problems, while hopefully maintaining the essential interesting challenges for cognition. The SeekWell microdomain developed by Steve Larson provides such a restriction for the problem of melodic expectation (Larson 1993). The microdomain involves the following features:

- The purpose of a problem in the domain is to produce an expected continuation to a sequence of notes
- The only material in a problem is a simple list of notes, where all notes are identical except for pitch (that is, dynamics, articulation, timbre, etc. are not involved)
- Notes are all of identical duration
- Notes begin at fixed metrical positions without any intervening rests
- Only one note sounds at once; the sound represented is monophonic
- Melodies in the domain are intended to be examples of Western tonal music

2.3.2 The Game

SeekWell also refers to a "game" that people can play. Playing the SeekWell game typically involves singing, playing, or presenting the score of a melody to a listener, who then sings, plays, or writes an expected continuation or series of continuations. The purpose of the game is simply to produce an answer to the question "what comes next?"

2.3.3 The Program

The goal of a SeekWell model is to simulate the typical results of a person (in particular, a listener of Western tonal music) playing the SeekWell game. Several models have been developed and implemented as computer programs to facilitate experimentation. The single-level and multi-level models have already been tested and compared with experimental results (Larson In Press), while the FARG model is the newest version, developed in this paper.

2.3.3.1 Single-Level Model

The single-level model simply involves predictions of melodic continuations using Larson's theory of musical forces. Although implemented as a computer program, it is straightforward (but tedious) to produce the model predictions by hand. Interestingly, the model first determines the possible pitch alphabets in use by examining the pitches present in the given melody. Next it examines the final two or three notes of the melody and produces gravity, magnetism, and inertia predictions based on these notes, ignoring the rest of the melody. These predictions are assigned probabilities according to the theory. Finally, if desired the resulting predictions can be fed back into the system as new input so that is can predict what will follow each of those possible continuations, generating a more extended prediction.

2.3.3.2 Multi-Level Model

The multi-level model is more sophisticated, at the cost of requiring up-front help from the user in the form of additional analysis. Instead of just inputting the melody, the user of the system must provide both the melody's key and a restricted form of a Schenkerian analysis of the melody. This analysis includes a series of levels ranging from the foreground through to the background. The notes at each level are related to the notes of the next-deeper level via an explicit description of the embellishment function. For example, a note might be labeled as an "upper-neighbor prefix" or a group of notes might be a "stepwise ascending suffix" to another tone.

The multi-level model makes predictions by applying the predictions of the single-level model to all levels of the analysis beginning with the deepest level. Of particular note is the use of the embellishment descriptions: a prediction based on inertia will indicate that the notes at that level will continue to go "the same way", where this is no longer an explicit direction (such as up, down, or stationary) but instead indicates that the embellishment pattern specified will continue for subsequent notes. Once the surface predictions have been made, the model presents these to the user as expected continuations.

2.3.3.3 FARG Model

The next logical step for a SeekWell model is to build on the success of the multi-level model by proposing a version in which the key and Schenkerian analysis do not need to be supplied to the model. Larson gives a set of instructions to provide a flavor of a complete version of the theory, without making it explicit. Of the six items he lists, four of the items reference the theory of musical forces and another describes how to apply the forces predictions to separate levels of the hierarchy. These items are already present to some extent in the multi-level model. However, the other item is more novel:

> Build up an internal representation of the cue (the "analysis") that includes the key, the mode, the meter, and

> a hierarchical representation of the embellishment functions and rhythmic attributes of each note or group of notes and the traces they leave. Evaluate the quality of that analysis in terms of its simplicity and order – a kind of confidence rating. – (Larson In Press)

This simultaneous building-up of the analysis is the critical component of the model I address in this paper. Larson goes on to present the idea for modeling all the parts of this step, along with the other five steps listed, in a manner consistent with other FARG work:

> While I have numbered these instructions, we should consider them as taking place in parallel, and influencing one another, until a potential completion is chosen.
>
> This set of instructions resembles computer models of analogous cognitive tasks (pattern finding, sequence extrapolation, and analogy making) created at the Center for research on Concepts and Cognition (Hofstadter et al, 1995). The success and sophistication of those programs suggests that a complete implementation of this set of instructions, while a very complex task, is possible and will likely by quite informative about music cognition. – (Larson In Press)

The remainder of this paper aims to expand on these ideas to propose a new model for SeekWell (I refer to this new model as simply “SeekWell” in the following sections).

3 Model Simulation by Thought-Experiment

3.1 Explanation

As a first step towards describing SeekWell, I wrote down some simple melodies and played the SeekWell game, recording the continuations I expected. I also wrote down some of my thought-processes involved during the listening and the completion phases of the game. Alternatively, I would record what I imagined the complete model or even the multi-level model going through for simple cases. The lessons learned from this introspection process were considered later while incorporating the FARG architecture into SeekWell.

3.2 Examples

Two examples of the game with my commentary follow. The first melody was suggested by Steve Larson, while I came up with the second.

To clarify my descriptions, I should point out that I had some trouble deciding what the appropriate hierarchical structure of each melody would be in each case. In particular, I didn’t have any melodies in which more than two levels of hierarchy were apparent. Thus, I usually refer to the “surface” when I mean all the notes of the melody but I say “background” when I’m thinking about the next level of structural importance. In some cases this should probably have been called a “middle level” with a true background level hidden deeper in the structure. I was also trying to conform to the multi-level model by requiring that each deeper level of structure retained notes in a regular metric fashion; these analyses are constrained metrically in a way that a normal

Schenkerian analysis is not. To see this, consider how in the first example below, every note at the deeper level is taken from notes on beat 1, and each measure contributes exactly one note to the deeper structure.

Finally, note that in the examples, the first line of music is the given cue melody, while the other lines of music provide the possible continuations along with implied analyses and key signatures.

3.2.1 Example 1

In listening to this melody, presented by Steve Larson, the key is established easily by the ascending scale through the tonic triad, with the emphasis on C and E. Here the collection of pitches involved, along with the primacy of the C pitch, give strong evidence that a naïve algorithm should be able to cope with.

While the music is being played, I assume the meter to be duple at first, but the presence of the repeated E, where repetition is not the norm in the rest of the cue, suggests that something interesting is happening. One possibility is that the second E falls on a strong beat, suggesting 3 beats per measure. This is not confirmed until later on, but after the first 5 beats triple meter is a strong working hypothesis.

Given the meter and key signature, the background structure can be extrapolated from the strong beat of each measure via a simple rule. In this case it gives C followed by E.

For this example, the key and meter determination was not very complicated, but now we turn to the problem of writing down an expectation for what follows.

3.2.1.1 Case 1: Staffs b (background) and c (surface)

Now, to compose the expected continuation, the multi-level model would examine the structure at level B. At this level, the C and E notes suggest a reference alphabet of the tonic triad, with the tonic and octave comprising the goal alphabet. Inertia and magnetism both point to continuing to G instead of giving in to gravity and falling back to C, so the next note of the prediction is G followed by the inertial continuation on to the upper C. This completes the background level completion.

At the surface level, the given rules determine that the major scale is the reference alphabet, while the tonic triad is the goal alphabet. To fill in the surface structure an analysis of the other notes besides the strong C and E is necessary. In both measures, the first note is followed by two ascending steps through the reference alphabet, so this pattern is recorded and used to complete measure 3 which is already known to start with G, producing the G-A-B continuation. Finally, in measure 4, we note that both the foreground and background levels have reached a goal pitch – C in both cases – so this is an appropriate place to end the expected continuation.

3.2.1.2 Case 2: Staffs d and e

A second expected continuation is also possible. Consider that after analyzing the given cue in the same manner as above the listener entertains the possibility of continuing on to G and then on to C in the background level as in staff b. However, due to the expectation for the 2nd and 3rd beats of each measure to continue rising by step, this background implies a melody that continues to ascend in the same direction for a whole octave. This is a rather long span to continue in the same direction without interruption, so pressure to turn around in the middle is introduced.

A simple hypothetical idea to “turn around” the melody would be to start measure 3 with a C, giving in to gravity at the background level shown in staff d. This idea might be workable, although the jump back down from G to C in staff E might sound out of place after the stepwise ascending beginning of the melody; this expectation of stepwise continuity at the surface introduces another pressure.

Instead of modifying the background level, another idea is to modify the surface structure in staff e. Retaining the G in the first beat and discarding the idea of continuing upwards to A and B, the initial analysis of beats 1 and 2 in the first two measures might be generalized to say “continue by stepping twice in the same direction for two beats”. With the constraint of going upwards thus removed, the continuation could proceed downwards through F back to E. Following this lead, the background structure might also “turn around” in measure 4, going back down to E just in time for the E to be repeated (following the pattern of the 1st beat of each measure repeating the 3rd beat of its preceding measure). Finally, the background could continue due to pressure from inertia and gravity back down to C, as would the surface structure, coming to rest on the goal of the initial C in measure 5.

This potential continuation, unfortunately, lacks the strength of the first one (staffs b-c) due to its mundane symmetry (with a repeated high note in the very middle of the phrase) and the implied tonic harmony throughout. The shift in direction in measure 3 weakens the background structure which was able to climb to the upper C in the first version. The first version also implied a shift of harmony to the dominant in measure 3, cadencing on I on the last note; no cadential pattern is in effect at the conclusion of this melody.

3.2.1.3 Case 3: Staffs f and g

Examining the weaknesses of the second continuation might lead to several different improvements. Two of these in particular seem promising as places for departure: the lack of harmonic changes and the lack of an implied cadential pattern.

Considering these ideas we can return to deciding how to continue the background structure in measure 3. Because the G has already been tried both times before, we again search for a new note. One of the pressures at the surface level – reversing the direction of the melody – is still in force here, so a note below G, such as E, seems in order. However, we also want to change the harmony, possibly to V in order to establish a cadence. V gives us the most likely choice of G again, which has already been used effectively in the first attempt; B and D are too far away from the G at the end of measure 2 to be useful while preserving stepwise continuity. This leads us to consider other chords such as V7. Indeed, the F in V7 seems promising, because it is just a step away from the melody’s G and also reverses direction as desired. (Staff f)

Several possibilities emerge here, such as continuing with F-G-A in measure 3 -- although this implies the harmony IV instead of V7 it seems reasonable. However, because we have reversed the direction of the melody at the background level it seems natural to reverse the pattern analyzed for beats 2 and 3 in the first measures as well. Where the previous pattern was “continue up for 2 steps” this would be effectively swapped for “continue down for 2 steps”, leading to the continuation “F-E-D”. This also has the virtue of implying the V7 intended (or perhaps ii or vii, of course).

Finally, to complete the phrase we need to continue the background level, because F is an unstable tone to stop on. The most likely candidate is E, predicted by giving in to gravity and magnetism at the background level. This also implies a tonic harmony in the

4th measure which makes sense following the dominant harmony in the previous measure. This continuation has reached a point of repose here at this imperfect cadence, so this is a good point to stop the continuation.

3.2.2 Example 2

The initial challenge provided by this example is determining the key. Because there is no obvious way to determine the key it makes sense to continue on to examine the implied meter, which may be expected to help shed light on the key problem. In the absence of other information we can safely assume duple meter. In this case, the 5th note stands out because of the leap to and away from the note, so it would make sense for it to be on a strong beat. Additionally, the pattern of notes in each measure (if we assume 4 beats per measure) is consistent; there is a pattern involving descending notes with identical 2nd and 3rd notes in each measure.

Returning to the question of key, our metric analysis has provided some useful insight. First, due to the collection of pitches present, we can assume C major and G major are the most likely choices. Next, if we just consider the strongest beats – the first beat of each of the two measures – we see a rising C-D movement. This might seem to be a typical way to begin a melody in C major but it seems out of place for G major. Looking at the notes another way, we might notice an overall descending line including C-B-A-G. This also sounds promising for C major.

The next question is one of determining the current embellishment pattern in order to predict the continuation in the following measures. The most obvious bit of pattern is the repeated 2nd note in each measure. Additionally, it is clear that each note after the first in each measure is lower than the previous. In m1 the pattern is stepwise but in m2 there is a leap down away from the first note, followed by the stepwise descent.

A sample continuation would take the background C-D line into account and write something like E-D-D-C for m3. For m4, the C-D-E line might keep rising up to F based on inertia, but leaping down in the C triad back to C also sounds plausible due to voice leading because the surface structure is a melodic line leading down to C (see staffs b-c). As the tonic, C is certainly a goal tone and this might be a phrase ending.

A slightly different continuation might give thought to harmony, and if it decided to retain tonic harmony then m3-4 might turn out the same as above. However, if we expected a half cadence on V then we might expect the background level to continue to a note of the V chord in m4. In this case, D seems plausible based on a gravity prediction from the E above and a magnetism prediction based on moving to V harmony. In hearing the continuation based on the background C-D-E-D line, we might also employ the device of m. 2 to leap down between the 1st and second beats, because this has the benefit of emphasizing C major, which wasn't clear enough in the beginning measures; this would lead to E-C-C for the first 3 beats of m3. The last note could continue down to B; however, there is a force expecting it to move up to D for improved voice-leading with the D expected in m4. This involves relaxing the pattern noticed in the first 2 measures by requiring only a stepwise motion to beat 4 instead of a step downwards. (see staffs d-e).

These two continuations notwithstanding, there is still another possibility, based on seeking an explanation for the pattern of repeated notes in the middle of each measure combined with the enticing C-B-A-G descending line hidden away in the opening 2 measures. The repeated notes on beats 2 and 3 sound too mechanical for some reason, but if we heard the second as an accented nonharmonic tone it might be more interesting. Indeed, if we shift our perception of the measures so that the 2nd note of each repeated pair falls on the downbeat, we hear the first two notes as a pickup measure and the C-B-A-G line is more prominent. Ignoring the pickup measure for the moment, we might also imagine the familiar idiom of a series of 4-3 melodic suspensions, implying a bass line such as F-E in measures 1 and 2, along with a possible analysis of ii6 and I6 in the first measures.

Considering the background structure, we now see a prominent C-B-A line if we include the C in the pickup measure (even though it's not in the same metric position as the B and A). This line has the obvious continuation down to G, so we can predict measure 3 to go something like A-G-C-G (see staffs f-g) where the jump up to C is expected by reexamining the embellishment pattern in the 1st measure and seeing that the 3rd note is a leap to and away from a harmonic tone. This continuation in m. 3 also helps solidify the feeling of C major.

Finally, m. 4 and 5 might drive towards a cadence on V to help clarify the key. The sample continuation is a bit hard to explain, although I think it's due to an expectation to keep the background descending line moving instead of stopping on G; it leaps down to the next note in the V chord, D. This inspires the voice leading in m. 4 where the 4th note is predicted to be E instead of F in order to move stepwise down to the final D. In addition, the 3rd note of that measure might be predicted to be B (the pattern here is a fourth above the 2nd note), but the desire to avoid the augmented fourth leap causes the B to be replaced with a smaller leap up to A. This also preserves the leap of a fourth downward from beat 3 to beat 4, as well as causing the leap to be to and away from a harmonic tone (as in the previous measures) if we analyze this measure as the ii chord.

4 Model Features

At this point we are prepared to offer a more detailed description of SeekWell. The architecture is described below after reviewing the properties we consider crucial to the model.

4.1 Desired Properties

4.1.1 Time-Dependant Presentation

Although not stated explicitly for the single-level or multi-level models, because SeekWell is intended as a cognitive model, it is important that it respects the perception of music as it unfolds through time. It is reasonable to expect that the perception of a melody we be different on repeated hearings, but for now we concentrate on the case where the listener is hearing a novel melody.

A concrete consequence of this is that we would not expect the model to be able to revise its analysis of music heard many measures earlier. Recent events can certainly be re-heard and re-analyzed, but as music is generally progressing forward through time, there is not an opportunity to make detailed model revisions. Indeed, assuming that musical events are stored in the brain as symbolic mental representations, there is no original event to return to for revision. That is, people don't generally have perfect memories for sound – a model shouldn't rely on a perfect memory for past events either.

4.1.2 Pressures: Top-Down and Bottom-Up

We expect that analysis decisions in SeekWell will rely on a combination of top-down and bottom-up pressures. That is, decisions made while "listening" to an input cue will be influenced by the low-level surface features of the music, such as expecting a leading tone to resolve to the tonic. Top-down decisions, on the other hand, relate to larger-scale concerns. These are provided by deeper levels of the analysis hierarchy in this model: the surface-level tendencies of melody might be overridden due to the need for the background line to move in a certain direction. Depending on context, various pressures might be more or less significant, so in some cases bottom-up processing will predominate while in others the top-down needs become more urgent than the surface features.

4.1.3 Mental "grouping"

The structure of the embellishment hierarchies generated require that notes are related to other notes in certain relationships. Several notes might form a larger group to be processed by subcognitive agents as a unit. Notes also might relate to others via an embellishment function. A model of this mental grouping should rely on Gestalt psychology ideas, or more colorfully, groups of notes may be built up using "bottom-up gestalt-based agglomeration heuristics" (Hofstadter and Fluid Analogies Research Group. 1995).

The mental representation of an embellishment hierarchy is built up over time as a piece is heard. However, the current and recent parts of the representation should shift fluidly in response to musical events. That is, even when notes are related together, these relationships can be broken if necessary due to other pressures to form different relationships.

4.2 Architecture

The proposed model for SeekWell follows the lead of other FARG models by involving several different interacting components. Although the names and ideas of the main components below are borrowed from other projects (Hofstadter and Fluid Analogies Research Group. 1995), each of those projects uses the components in a domain-specific manner. In fact, surprisingly little has been reused between FARG projects in the past. Likewise, here the big picture may look similar but the application to the SeekWell domain is unique.

4.2.1 Workspace (Embellishment Hierarchy)

The Workspace represents the internal cognitive space where "thinking" happens. The actual analysis and expected continuations are generated here. The workspace is intended to convey an image of bustling construction area: subcognitive agents are busy working in the workspace to construct an analysis of notes that have been heard in the form of an embellishment hierarchy. The working agents might build up structures such as groups of notes or analytical links between notes, or tear down those structures to try something different. The particular types of data stored and worked on in the workspace follow:

4.2.1.1 Surface-Level Notes

As new notes are presented to the model, they are stored in the workspace at the surface level of the growing embellishment hierarchy. These notes are marked in such a way that the system can tell that they were actually heard, as opposed to notes that have been predicted or that are inferred at a deeper level of the hierarchy. As time goes on, this marking fades until the system can not tell the difference between heard notes and other notes.

4.2.1.2 Key Area

For each segment of the surface level, a working theory about the key of that section is maintained. If the key is ambiguous (for example, at the very start of a melody) then multiple key ideas can be maintained, although each potential key has a likelihood rating associated with it. Agents work to clarify the understood key and to ensure that the once a key has been decided, it is relatively stable; it takes a lot of pressure to change the key once firmly established. Whenever the working hypothesis about the key area changes, action is initiated in the workspace in order to bring the analysis in line with the new key.

4.2.1.3 Meter

The implied meter for each segment of music is recorded in much the same manner as the key. Again, at the start of a piece the meter may be particularly ambiguous but agents work to understand the meter as the melody is presented. If the understood meter changes, the analysis must be updated immediately to deal with the new meter.

4.2.1.4 Embellishment Hierarchy

The embellishment hierarchy is the most significant part of the Workspace. The hierarchy consists of several levels of notes, where at each level some notes are marked as being “embellishing” while others are marked as “structural” for that level. Embellishing notes are not present in the next level “up” (by which we mean the level further in the “background”. Structural notes are present in the next level up, but at that level they may be either structural or embellishing.

As noted earlier, the upper levels are metrically fixed relative to the surface level. That is, at a certain level the first beat of each measure might be structural, while all other notes would be embellishing. Thus the hierarchy is a restricted type of Schenkerian analysis. This is considered to be a restriction of the microdomain of SeekWell, but it may be necessary to lift the metric constraint if it causes unexpected results.

To enforce the constraint that the musical input is processed through time, the hierarchy automatically “hardens” while time passes after being built. For example, an analysis from 8 notes in the past would take much more pressure to “break” and change than the analysis for a short time interval containing the most recently heard note. An interesting effect of this is that the resulting analysis might not have the global-scale logic and coherence of a typical theoretical analysis. This is intentional: the analysis should reflect the piece as-perceived during an initial hearing, rather than as-analyzed after-the-fact.

Note that the use of a “construction area” metaphor here is intentional; the image of workers building a cognitive structure is evocative of a possible way that thought patterns emerge and form into structures.

4.2.2 WorkerCollection

The WorkerCollection maintains a group of workers prepared to do some work in the workspace. Each worker (I sometimes use the term “agent”) has its own urgency and its own typical task to perform. The WorkerCollection is called the ”Coderack” in other

FARG models, but I find it confusing to blur the distinction between the theoretical model and the computer implementation by using the computer term “code”. Likewise, I have used the term “worker” in lieu of “codelet”. Workers are individual procedures that carry out small tasks. At any time, many workers are available in the WorkerCollection, but they do not perform any tasks until selected by the controller. Workers are selected stochastically based on their associated urgency values. Several workers may work in the workspace at once; the system simulates this parallel activity. The urgency value also affects the speed at which a worker proceeds; in the simulation, a worker with high urgency is allowed more time to do work than other workers of lower urgency. Workers are also able to add new workers to the collection if necessary for a given task.

Crucial to the model is the selection of the types of workers and the definition of what they will be capable of. Some workers are intended to focus on bottom-up tasks such as analyzing individual notes, while others deal with top-down ideas such as considering high levels of the hierarchy or big-picture concepts as stored in the Conceptual Memory. The question of exactly what worker types to include demands more research, but I propose a short sample list of useful worker types:

- Describer: Describe an embellishment (embellishment types are detailed below)
- Generalizer: Generalizes an existing embellishment description
- Specifier: Makes an embellishment description more specific
- Grouper: Group two or more notes
- Breaker: Destroy a group of notes
- Structuralist: Reanalyze an embellishment note into a structural note
- Post-Structuralist: Remove structural note analyses
- Rhythm-finder: Look for notes demarcating metric boundaries
- Key-finder: Looks for evidence pointing to the key of the melody
- Predictor (three types: gravity, inertia, and magnetism): Makes musical force-based predictions at any level of the hierarchy
- Resolver: Notices differences between expectations and predictions, and tries to fix the problem.
- Sequencer: looks for sequential patterns
- Trace-finder: looks for auralized traces, left by a melody leaping away from an unstable note
- Line-finder: groups disjoint notes together based on considerations of auralized traces, to find instances of lines in a compound melody
- Etc.

A critical worker is the Describer, because it creates the embellishment analysis used in inertia predictions. The following sample list gives some of the embellishment types that the Describer can attach to notes. Many of them can be used in combination, unless they are mutually exclusive. For example, a note can be described as a neighbor of another note, a prefix neighbor, a suffix neighbor, or even a suffix upper chromatic neighbor:

- Prefix/suffix
- Neighbor
- Leap/Step
- Chromatic/diatonic
- Upper/Lower
- Major/minor/augmented/diminished
- Unison/second/third/fourth/fifth/sixth/seventh/octave/ninth/etc.
- Passing tone/escape tone/appoggiatura/suspension
- "Group" (this refers to a whole group of notes that serves an embellishing function; the individual notes in the group are also described with this list of terms)
- Pedal
- Etc.

Each of the worker types described earlier will function differently based on the contents of the Conceptual Memory, described in the next section. For example, if the concept of "sequence" is activated, workers may be inclined to search carefully for patterns of notes that indicate sequencing. If a concept such as "stepwise motion" is activated, Describer workers will be more likely to label embellishing notes based on stepwise connectivity than based on actual interval qualities. Similarly, a Describer might choose more general embellishing descriptions from the list above when some concepts are active, but other concepts might indicate the need for precision in embellishing description, so that more specific terms would be used in combination for each description.

Also note the important rule of the "Breaker", "Generalizer", and "Specifier". These workers modify the existing analysis to provide the desired fluidity in the model. If system pressures indicate the analysis is not functioning well (due to divergence between expectations and inputs or an overly complex analysis) these workers can be called up to start changing things in hopes of improving the situation. There is no guarantee that random breaking of groups will help, but it allows the possibility of a better analysis being re-formed.

4.2.3 Conceptual Memory

The conceptual memory is a fixed representation of musical knowledge, such as the knowledge one maintains about Western tonal music after much experience. Examples of such knowledge are: the difference between large and small intervals, between stepwise motion and leaps, the desire for the buildup to a melodic climax on high notes, understanding of sequence, known pitch alphabets such as the major and minor scales, and simple harmonic schemes such as V->I. Each of these concepts is represented by a node in the conceptual memory. Concepts can be linked to each other with connections of varying strengths; the conceptual memory looks like a connectionist system but its purpose is simply to represent permanent knowledge and to provide focus on certain pieces of knowledge; it doesn't learn new concepts, for instance. Each concept can be become more or less activated due to the context of the workspace. When a concept becomes more activated, it will also spread some of this activation to nodes linked to it, so that focus will be given to several related concepts at once. When concepts are activated, certain workers may behave differently, in manners appropriate to the current concepts. Also, different workers may be placed in the WorkerCollection as concept activation increases.

The conceptual memory provides an important top-down influence on the processes in the workspace: activated concepts are rather persistent in contrast to the ephemeral nature of the individual workers. An activated concept will tend to direct the progress towards a particular type of analysis, even as surface-level features are able to have a small influence on the activated concepts. Concept "slippage" can also occur in the conceptual memory; a related but distinct concept might overtake an existing activated concept, causing the top-down focus of the model to change significantly to reflect the new concept.

4.2.4 Metacontroller

The Metacontroller is the component that integrates the various parts of the system and controls the process of running the system. It also maintains record of the system "temperature", which indicates how satisfied the system is with its current state. The most prominent feature of the Metacontroller is a main loop for "listening"; within this loop the following events happen:

- The next note of the cue is added to the Workspace
- Older structures in the Workspace are hardened to make them more permanent
- The WorkerCollection is sends some workers to the Workspace
- Activations in the ConceptualMemory shift in response to actions in the Workspace
- The workers are allowed to work until the system temperature becomes low
- The cycle repeats until all notes in the cue have been presented

Next, the system enters a stage where it generates expectations; this looks almost like the main loop above:

- The next expected note is added to the Workspace as a prediction
- Older structures in the Workspace are hardened to make them more permanent
- The WorkerCollection sends some workers to the Workspace
- Activations in the ConceptualMemory shift in response to actions in the Workspace
- The workers are allowed to work until the system temperature becomes low
- The cycle repeats until at every level, the expectation is static because a point of closure has been reached

Note that these two stages are barely distinguishable, except that in the first stage the system is listening to new notes from the cue while in the second stage it is listening to new notes generated by its own expectations. Indeed, in the "listening" stage the workers also generate expectations; the difference is that these expectations may be overruled by actual notes in the input. This provides a powerful force to guide the workers, in fact: when the actual input is not in-line with the expectation, this should trigger additional work to fit the working embellishment hierarchy to the actual new note.

After both loops finish, the final expectation along with the generated analysis is presented to the user.

4.3 Application to Problems

Now that the main components of the model have been described, we discuss a few interesting problems and show how the model may be able to cope with the cognitive issues involved.

4.3.1 How Does the Model Determine Key, Analysis, and Meter?

When the listening stage begins, the key is unknown, so workers with high urgency are tasked with determining the key. In parallel, other workers are building the embellishment hierarchy and associated metric information. The key-finder workers can use the hierarchy to great advantage, because structurally important notes will have more influence on determining possible keys. The actual algorithms used by the key-finders will be based on a simplified version of Temperley's algorithm to determine key (Temperley 2001). However, the algorithm should not need to work perfectly to get a good result: by combining possible key ideas with the growing embellishment hierarchy, the most likely key should emerge naturally.

Conversely, the workers building up the hierarchy and metric structure will need to know the working key, so they turn to the key-finders for the needed information. As ideas of the current key change, the hierarchy will necessarily go through major changes, but our expectation is that gradually a stable picture of both key and hierarchy will emerge; when this stability emerges the system can be more sure that the key and analysis are both reasonable. As processing continues and more notes of the melody are input, they will either further confirm and solidify the key choice or present problems that indicate the key was wrong. If too much time has passed and the earlier analysis has "hardened" significantly, the new key area may signify a modulation instead of an incorrect earlier analysis.

Of course, this is just a hypothesis; it will be interesting to see if this process works as described on an implemented system.

4.3.2 Expectation and Musical Forces

As mentioned above, expected continuations are being generated constantly by Predictor workers; these expectations are generated on all levels of the hierarchy. However, in contrast to the multi-level model, the background-level expectations do not necessarily come before those of the surface-level. Instead, the expectations at adjacent levels will be considered by the predictors. If a prediction based on musical forces is in conflict with another level, the other level might overcome the tendency of the musical force. In this context, the musical forces simply provide another source of pressure that influences decisions in the system. Musical forces at multiple levels can conflict with each other; out of this complexity, we hope that dynamic but realistic predictions may emerge.

Eric Isaacson has raised the question of whether certain hierarchical levels are more likely to prevail in such a conflict. For example, perhaps surface features pressures are stronger than those of a middleground level. This question merits careful consideration, as does the generalized question of how to "weight" different interacting pressures, as discussed in the following Challenges section.

5 Challenges

The implementation and use of the proposed architecture may result in several challenges. Several of the issues are listed here.

5.1 Complex Domain

Even though SeekWell is a restricted microdomain, there is still a great deal of complexity due to factors such as musical rhythm, implied harmony, and compound melody. It takes much domain-specific information to encode the knowledge of the workers and in the Conceptual Network.

5.2 Parameter Values

A system with many components such as this one faces the problem of how the different components interact meaningfully. In particular, when there are arbitrary units involved such as the "activation" of a concept in the Conceptual Network or the "urgency" of a particular worker, how do we guarantee that these values can be set in a mathematically appropriate way to allow interesting behavior? One can easily imagine that if certain parameter values are not "lined up" correctly, a highly activated conceptual network node may continue adding certain workers to the WorkerCollection with high urgency so that little happens except the repeated work of these workers. At some point during implementation there will be questions such as "when a node in the Conceptual Network is activated with value x, what activation is sent on to its neighbor node?" There are no principled ways to answer these questions; instead, the model will involve many quite arbitrary decisions. This is a problem, I believe, with many simulation systems and models, and should be addressed in the future.

5.3 Harmony

Although Dr. Larson would like to see SeekWell working without any explicit knowledge of harmonic functions, I believe it may be difficult to have the model behave in the same way as a person who is used to thinking in terms of tonal harmony. It would be an interesting theoretic question, however, to see how much harmonic knowledge is unnecessary when an embellishment hierarchy is being used and attention is focused on voice-leading concerns. Perhaps melodic expectation will turn out to be a useful domain in which to test ideas of Schenker and other theorists.

5.4 Flexibility of the Hierarchical Analysis

As mentioned earlier, the metrically-fixed way in which levels in the hierarchy are connected to higher levels may result in too-rigid processing. Normally expected melody continuations, for instance, might be ignored due to a simple extra beat in a phrase that confounds the structure of the hierarchy. If cases are demonstrated where the hierarchy is too inflexible it may be necessary to add metric flexibility, but that would surely increase the system complexity as well.

6 Future Work

An obvious next step for SeekWell is the complete formulation of this model including very specific information about the concepts to include in the Conceptual Memory, the precise description of a whole population of workers, and other details that prevent immediate implementation as a computer system. It would also be useful to simulate a very simple scenario for initial key detection to show its dependence on the generation of the embellishment hierarchy.

The next step, then, is to implement the model as a computer program, as described in the previous section. Finally, a working version of the model could be used to test the hypotheses of the system in more detail and to compare its results with human studies such as those already performed at the University of Oregon.

A related project would be to use the SeekWell model as part of a model for melodic error detection. Hearing a note as “wrong” is in some sense the opposite of hearing a note as the expected continuation, so SeekWell might provide some insight to this related problem.

A more distant future goal is a more formal characterization of FARG architectures such as this one. While this type of architecture is compelling to me, it is frustrating that its use is so ad-hoc. That is, the best way to examine this model seems to be to implement it and test it; it would be preferable to analyze the formal nature of the model on paper without the necessity of simulation. However, this may be a mathematician’s pipe-dream; some systems may be complex enough that simulation is the best analysis possible.

7 Bibliography

Hofstadter, D. R. and Fluid Analogies Research Group. (1995). Fluid concepts & creative analogies : computer models of the fundamental mechanisms of thought. New York, Basic Books.

Larson, S. (1993). Seek Well: A Creative Microdomain for Studying Expressive Meaning in Music, Available (from Center for Research on Concepts and Cognition; 510 North Fess; Bloomington, IN 47408).

Larson, S. (In Press). " Musical Forces and Melodic Expectations: Comparing Computer Models with Experimental Results." Music Perception.

Meredith, M. J. (1986). Seek-Whence: a model of pattern perception**:** xiii, 189 leaves.

Minsky, M. L. (1988). The society of mind. New York, Simon & Schuster.

Temperley, D. (2001). The cognition of basic musical structures. Cambridge, Mass., MIT Press.