Music Informatics

Alan Smaill

Mar 29 2018

Today

- Human/machine interaction in performance
- Improvisation and imitation
- the OMax project
- examinable material

Musical interaction in performance

There are many ways that artificial (virtual) musicians can get involved in musical performance with human musicians.

- Play alternately, one after another, following score, with obvious cues to alternate
- Play concurrently following score, coordination by score following algorithm
- ...
- Improvise collectively, without distinction between human and virtual performers.

These get progressively harder to achieve. Still, there has been a lot of work on machine improvisation.

See for example Robert Rowe, Interactive Music Systems, MIT Press, 1992 for some idea of what was going on at that time.

More recently there has been impressive work incorporating learning aspects that allow virtual agents to join in improvisation with very little experience of the sort of music they are improvising along with, while at the same time sounding believable musically. How is this done?

A recent approach by Pachet looks for probable continuations on the basis of music heard; prediction probabilities are influenced by a fitness function that is used to take account of the local musical context.

This approach does not attempt to plan ahead, or to adapt its behaviour in action.

See https://www.francoispachet.fr/continuator/

`http://francoispachet.fr/continuator/`

The OMax project at IRCAM has been researching into this problem. this paper on Anticipatory Machine Improvisation and Style Imitation is recommended:

http://www.springerlink.com/index/n63260485503v3g5.pdf

(This is accessible if you are logged in to EASE)

The slides follow the paper closely. The paper describes some aspects of the OMax systems;
clearly a lot of work has gone into supporting the level of interaction that can be heard in the demos.

Aspects of the approach

One traditional view of musical experience is that the role of expectation is very important:
the listener forms expectations about what is about to happen in the music, and the expectations may be fulfilled or not.

Here is used rather a notion of anticipatory system:

> *a system containing a predictive model of itself and/or its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a later instant*
>
> *Robert Rosen, Anticipatory Systems*

Other aspects

- planning the musical material
- multiple representations of music signals
- reinforcement learning
- respect for cognitive aspects
- multiple agents, each modelling different aspects of the music, each learning and suggesting how the music should proceed.

The terminology of agents has become prevalent recently;
note that it can be used to describe a complex entity, like a virtual musician, or simpler agents that combine together to achieve complex behaviour, while individually performing simple tasks.

Minsky's "The Society of Mind" (1985) is a readable collection of short notes on his view of AI and mind. It influenced later work on agents. He argues that this is the right way to approach an understanding of mind. In "The Society of Mind":

> *I'll call "Society of Mind" this scheme in which each mind is made of many smaller processes. These we'll call agents. Each mental agent by itself can only do some simple thing that needs no mind or no thought at all. Yet when we join these agents into societies—in certain very special ways—this leads to true intelligence.*

Minsky suggests that even simple tasks involve several agents.

> *What kinds of smaller entities cooperate inside your mind to do your work? To start to see how minds are like societies, try this: pick up a cup of tea!*
>
> *Your GRASPING agents want to keep hold of the cup.*
> *Your BALANCING agents want to keep the tea from spilling out.*
> *Your THIRST agents want you to drink the tea.*
> *Your MOVING agents want to get the cup to your lips.*

Notice the attribution of mental attribute to the agents (they want something).

Properties of Agents

Typically agents are expected to have the following properties:

- *autonomy*: the agent can evolve on its own, without being directly controlled from outside.
- *social interaction*: agents usually interact with other agents, sometimes in cooperation, and sometimes in competition.
- *reaction*: a *reactive agent* is one that takes account of its environment, and responds to changes in the environment.
- *goal-directed*: the agent has its own goals, and takes initiatives in order to meet these goals.

A system based on agents is well suited to a distributed implementation.

Prediction and Interaction

We have seen how methods such as Markov models can be used to predict how music will proceed, give a local context. This does not deal with the sorts of interactions between musicians that occur in improvisation – they can provoke quick changes of direction in the music, rather than the "expected" continuation.

Here multiple representations are important, to allow quick access to relevant facets of the musical process; on the ither hand, but systems using them usually need a lot of time to learn a given style.

The work also takes into account cognitive aspects of musical expectation, in particular on musical perception and memory.
On expectation, the suggestion is that there are competing hypotheses represented in the brain when listening to music (so the multiple agent view has some cognitive plausibility).
On memory, the different roles of short-term and long-term memory also influence the design of the system.

Anticipating via reinforcement learning (RL)

Idea of RL:
an agent is in some environment, and can take actions which can change the agent and the environment; it has a statistical policy of what actions to take in what situations.

- the agent tries to maximise the rewards for its action
- the policy is adapted depending on how successful each action is

Here the agents have an internal model of the environment, based on memory of input sequences that have been experienced. The rewards work as guides to encourage or discourage the use of parts of the learnt model.
The aim is to learn good ways to plan ahead for musical output.

Anticipation

The system can work in two ways:

- in interaction with musician(s);
- in self-listening mode

A difference is that the rewards in the first case encourage imitation, and in the second case discourage imitation, which would lead to boring repetitive music.

The representations are based on with MIDI signals.
From the input, the system constructs a number of different finite state automata that describe evolution of different aspects of the music, such as "derivative of pitch" (difference between successive pitches).

These automata are called factor oracles; they are designed to make the observed patterns kept efficiently, both for space (number of nodes), and time, in that the information needed to react in real time is accessible very easily.

Multiple Representations

The system can be configured with different choices of representation, eg:

- `pitch` as MIDI number
- `harmonic interval` as harmony relative to fixed pitch
- `duration` as multiples of lowest level pulse

and also the "derivatives" of these (difference between successive values in first two cases, quotient for duration).

It can also analyse intensity, articulation, ...

The rewards are designed to get better and better models of the music involved; want to have both

- state anticipation: what are the likely states the music is heading for?
- payoff anticipation: how get the best payoff when self-listening

More information in the paper —
you are not expected to follow the details of how this is achieved.

Why does this work?

Some features that enable this approach to succeed in real time:

- Choice of MIDI input:
  immediate accurate access to note timings, pitch, intensity
- Choice of representation:
  need to have the most useful information easily accessible; factor oracles allow compact analysis of the patterns that are analysed
- Getting the rewards right in the RL phase

See the OMax web site

`http://repmus.ircam.fr/omax/`

– there are some informative clips to listen to.

Summary

- Machine collaboration in musical improvisation
- Multiple representations, reinforcement learning approach
- OMax system and web site

Examinable material

- Basics of different representations (midi, WTM score, mp3, sound-file):
  strengths and weaknesses, what they make easy and what they make hard
- Converting between representations (midi <-> mp3/wav), what's involved (without technical details)
- WTM metrical hierarchy, what it is.
- Longuet-Higgins metrical analysis algorithm
  main idea, know some of the rules (not all)
- musical ambiguity, how recognise? how deal with by machine?
- 2-d pitch array, how use to recognise key and note spelling

Examinable material ctd

- Beat-tracking: what the problem is, Dixon's algorithm (in outline)
- Score-following: what the problem is; outline HMM approach; the main architecture of Raphael's system
- Musical grammars, what they look like, what they can be used for
- GTTM: what it tried to do, the main components, cognitive claim
- GTTM grouping rules: some examples of the rules; what choices need to be made to implement these rules.
- Rule-based systems: what they are, what needs to be put in, how control the use of the rules

Examinable material ctd

- paradigmatic analysis: what it is, algorithm for carrying it out (parametrised on notion of similarity)
- Cope's style imitation: how he did it, notion of signature
- Xenakis: how generate music using statistical ideas, how generate musical curves from straight lines.
  UPIC and successor system – human interface to the system.
- Components of improvisation system in OMax
- Learning at different levels; what is a probabilistic grammar