

# Reconcile art and culture on the Web : Lessen the importance of *instantiation* so creation can better *fiction*

Francis Rousseaux<sup>1</sup>, Alain Bonardi<sup>2</sup>

<sup>1</sup> Laboratoire d'Etudes et de Recherches en Informatique  
University of Reims Champagne Ardennes (URCA)  
IST SemanticHIFI European project team  
Institut de Recherche et de Coordination Acoustique/Musique (IRCAM)  
1, place Igor-Stravinsky  
75004 Paris France  
francis.rousseau@ircam.fr

<sup>2</sup> Artificial Intelligence and Mobile Robotics Research Team  
University of Paris 8  
Maison des Sciences de l'Homme Paris-Nord  
4, rue de la Croix-Faron  
93210 La Plaine Saint-Denis France  
alain.bonardi@wanadoo.fr

**Abstract.** In this paper, we show how the practice of buying CDs surreptitiously conditions our musical activities, and how the slowly-evolving practice of classifying structured by buying and the notion of genre, will disappear, as develops an *ad-hoc* organization of auditory samples, centered on prototypes and similarity. Thus appears similarity-based calculus, acting on considerable masses of samples, which continuously change the balance of the man-machine dialogue, in an appeal started again and again to compare information. We also show how some artistic creations proceed in the same way. Art and culture are linked in processes that take advantage of their massive digitalization.

## 1 Preliminary Notes

*Instantiation* is often used by computer scientists, which comes from the word *instance*, which means *example*, *case*; instantiation somehow generalizes the operation used by mathematicians by which a numerical value is assigned to a variable. To speak about reality, computer scientists instantiate abstract categories, thus decreeing that this or that entity is a specific instantiation of a category, which itself is linked to other categories by general hierarchies and/or formal properties. The whole device [10] is sometimes called an *ontology* —ontologies are assumed to describe sections of mundane knowledge widely used in artificial intelligence, sometimes an *object-oriented design*.

The neologism *to fiction* is used here to remind us that an interactive computer device based on symbol manipulation *works* only if the two following conditions are met: on the

one hand, that programs starting an execution have a correct syntax, and on the other hand, that the user agrees to play the interaction game with the device, thus recognizing it as an operational fiction.

## **2 Navigating in oceans of music: a fascinating indexing problem**

Using a classifying mode always implies more assumptions than it seems. And this is why the issue of classifying is an old philosophical issue [1], [11].

### **2.1 Today's craze for on-line music distribution**

If online music distribution has caused a lot of ink to flow lately [13], it is because three circumstances allow social, cultural and business demands to persist and crystallize, and even to be so important as to become strategic industrial necessities:

- people who buy CDs do not understand that the Web is still deaf to their appetite for music "à la carte" [12];
- music designers fight to escape the normative constraints implied by CDs;
- music producers and music companies are organized in a small number of multinational companies in competition (the famous *Majors*), and they fear that a major innovation will disrupt their balance.

We will limit our study to the investigation of the *indexing of musical objects*, that is to say the study of the organization that would enable accessing and listening to music in a differentiated way, through the Web as a matter of course.

### **2.2 CDs surreptitiously condition our musical activities**

To begin, let's consider the world of CDs: they are concrete materials which contain sequences of musical pieces, and buying these concrete objects on the market is a necessary step to access their content.

To be able to sell CDs, it is necessary to arrange them on displays, and to label the shelves and the trays. It is not obvious at all, but the decisions taken here—very often implicitly—on the way CDs are organized, will have a considerable influence on the way our musical activities are established and organized, and especially on music listening, which is supposed to be so ineffable!

In music stores, each tray is labeled to allow a meaningful organization of products, by making a compromise between the physical constraints of the store, the clients buying habits, and the label readability.

But the compromises made by department managers of record dealers rival each other in cleverness and originality—this is how some items are sometimes present in *several different trays* at once. Faced with such compromises, the promoters of the science of classification would often tear out their hair. But business efficiency pre-

vails in this case over scientific rigor, and systematic coherence matters less than productivity [8]!

It is indeed the *buying* activity which conditions the indexing mode of CDs on the shelves of CD stores. When a buyer recommends a CD to a friend who will potentially buy it, he will explicitly use the label categories, which semiotic system he has learnt without knowing it, by walking around wholesale dealers...And the labeling device soon becomes part of a cultural heritage; it won't be long before it is used to describe the whole range of musical activities, including music listening itself, affected even in its most meditative modality [16].

## 2.2 From CDs to digital sound files

Some think that forcing descriptive labels into digitalized online distributed *sound files*, is a good way to elaborate online music distribution services. These labels then become what is called *meta-information*.

This solution is not devoid of interest for the comfort and the cognitive ergonomics of the users of online services, because it allows them to question machines in the way they questioned labels not long ago. Walking around spaces cluttered with shelves is here simulated by using pop up menus and lists on a screen. New problems will certainly arise when downloaded files will be organized, essentially from file names [9].

More generally, those who promote these solutions are not unaware of their built-in drawbacks: they know that the price to pay to maintain and update this meta-information is high, that chances are weak that we will be able to extract them automatically from musical data, and that the labeling system itself must first be rationalized to become compatible with a computerized processing.

But they often fail to understand the most reactionary aspect of their viewpoint. If we follow Gilbert Simondon and his theoretical proposal on the existence mode of technical objects [14], when mechanisms of online music distribution will come into being, they will inexorably be far from these inevitably temporary solutions ...

Drawing our inspiration from Simondon, we can hypothesize that with the loss of interest for CDs and the end of the requirement to buy the medium *before* listening to music, we will witness the corollary fall of an operational fiction, that was simplistic, but effective nonetheless. The object of the fantasy will not be to access the CD anymore, but the piece of music, or the sample, in a way that we must clarify. Because other forms of musical activities will emerge to make possible the organization of musical objects, which will be more often dedicated and signed. Thus the normative activity of CD buying will give way to situated actions, that will be part of more and more differentiated and singular projects.

## 2.3 The object/activity/description triad

This is what we meant when we wrote about "music-ripping" [3], to mean that listening means signed listening/composing/producing, based on *samples*.

What is the array of listening situations? Will it be necessary to try to define it by referring to professional or amateur practices—the way a composer listens when he is trying to compose, the way anyone listens to music in the shower in the morning, etc. ? No, because these stereotypes pertain too much to the ancient era and say nothing about the world of listening as signed [5], that is to say about listening when it causes appropriations and mediations (digital ones are what matter to us).

Therefore, other typologies of signed listening situations will emerge shortly, in so far as they will be facilitated by the emerging technical systems of indexation and navigation [17] which will stimulate us to imagine them. The slowly-evolving practice of classifying *a priori*, structured by buying and the notion of *genre*, will disappear, as develops an appropriate organization of auditory samples, centered on *prototypes* and *similarity*, situated in the singularity of everyday practice.

### **3. Using calculation though staying the master of the game: lessen the importance of instantiation**

How is it possible to inspect a singularity in normative category systems and calculation procedures? These devices are not arbitrary—they have their own life, their individual and concrete nature, but the *relationship* between the singular and the particular linked to generalities could very well look arbitrary. In computer science, this is what *instantiation* is, though it has never been described, and its effective reduction probably implies, as an *infinite remedial task*, consequences on the systems of categories themselves and their individuation.

To be able to make sense of the effectiveness of calculation, singularities of the living world must have been *instantiated* with symbolic conceptual particulars, which will be enlisted in conceptual systems (semantic networks or other ontologies) meant to represent knowledge relationships and to make their approximation possible by calculable models.

If it is reasonable to consider these conceptual devices as technical abstract entities, which as such will possibly materialize (in Simondon's sense), and become more refined —also by integrating more and more sophisticated meta-models, it is probably unreasonable not to see that their materialization is determined by the gap between 1° what they pretend to represent in order to make calculation possible (finalized vision) and 2° what they tend to recapture from the roughness of the instantiation operation (original vision).

#### **3.1. Instantiation as a calculation *incitement***

Instantiation is indeed the part of computer science that has not been reflected upon. Better: instantiation is the fundamental usurpation of computer science, its radical cheating, its dead angle. This has serious consequences on thinking and research in computer science.

First, the spontaneous assimilation of singularities to particulars linked to devices of conceptual meshing leaves the whole task of representing and simulating reality to these devices. This is how research in computer science wears itself out setting up devices, forgetting from the start that their specifications are prescribed underground by the attempt to compensate for the obscurity of the operations of instantiation. This probably explains the infinite development of research on the representation of knowledge and ontologies, which aims at materializing (by category differentiation) devices that are supposed to be able to give light to a black hole.

Second, the consequence of the practice of computer scientists is often to lessen the importance of the operations of instantiation, as if to veil its un-reflected nature. Even if it is rare for a computer scientist to claim this inspiration explicitly—Frédéric Drouillon [6] however places it at the heart of his creative work, by assuming that is possible to shy away from instantiation by stealing already instantiated systems to enroll them in more complex systems, it often operates as background to research [15].

Among the implicit means that enable designers of computing systems to lessen the importance of instantiations, there is one that relies on a heuristic meaning of calculation and on a vision of the request-calculus interactivity which is on the edge of contradicting Turing's hypotheses on the investment of the machines' minds by dialog, which is at the source of their reputation of intelligence. In the context of searching for digital content using similarity, this means is usually used.

Because when a calculation is used in order to organize a great number of digital entities according to their similarity to a specific example, the symbolic interpretation of the calculation and of the symbols that are used for it must be given up, and the only possible outcome is its heuristic efficiency in selecting particulars, proposed as so many candidates which undergo a singular election to be accepted by a user. This is how the fundamentally heuristic nature of the resulting similarity-based calculus is expressed—for any singular demand, there is always a corresponding instance among the many particular candidates.

In this case, the multi-criteria instantiation is done conforming to a distance, which allows a strong semantic interpretation even less as it relies on different means of statistical comparison, which retain the coherent and reproducible behavior of a user, or of a community of users. The important gesture is still the choice of the user when he recruits a singularity among a pre-selection of particulars. And it is only in this way that calculation is interpreted, in an "interested" way.

It is necessary to add that the user does not have to make a definite choice, but that on the contrary, he is encouraged to perform again his gesture of heuristic enquiry by similarity, starting with a new singularity and if necessary with new specifications criteria of the desired distance. Therefore, in repeating the gesture of aided selection, an evanescence of instantiation continues, sliding from selection to selection. And the dialog with the calculation is not part of Turing's intuitions, somewhat similarly to how the balance of a "balanced" chemical reaction can always be displaced by withdrawing progressively the material resulting from one of the two ways of the reaction. Balance is meta-stable. The other side of calculation is of no interest to the user who does not try to build its model, but only to profit from it dynamically.

### **3.2 Heuristic and meta-stable similarity: a growing mobilization of calculation**

Let's come back to the example of online music distribution. The activity of buying CDs used to condition the indexation system. The activities around searches for on-line data will focus on ad hoc nuclei and idiosyncratic goals, which will closely unite reception and action,, and thus determine their own objects, which we call samples [3].

The *a priori* indexation is gradually replaced by a dynamic indexation, and its central paradigm is usually similarity. This implies a need for algorithms to calculate descriptors on the basis of digital contents themselves, and an attempt to match these (self-extracting) descriptors to appropriate categories in the context of committed *ad hoc* activities. Indexation devices can then be composed of meta-information, but also of labels calculated from musical contents. The MPEG7 and MPEG21 norms thus propose to put on the market descriptors created from automatic extraction procedures and names that identify descriptive qualities.

## **4. Drama Interlude: two ways of using instantiation in the inter-media play *La traversée de la nuit - Through the night***

The theater play<sup>1</sup> *La traversée de la nuit* by Geneviève de Gaulle [7] is the story of the author, imprisoned in the dungeon of the concentration camp in Ravensbrück.

### **4.1 Instantiating with neural networks the expression of a comedian on stage**

The work on memory is here linked to a neural network, which is used to recognize the emotional states from the voice of the comedian who says the whole text. Neural networks are a clever and fashionable way to conceal instantiation.

For a few months, the comedian Valérie Le Louédec declaimed her text in front of the computer, in preset moods: joy, sadness, resignation, anger, etc. After choosing a mood, the comedian tries to say the text in that state, no matter what the meaning of the text implies. For every pronounced sentence, a vector of twelve components is extracted from her voice: four for the vowels, four for the consonants, four for the prosody. Then these vectors need to be linked to the intended emotional state, and this is the learning phase for the neural network.

Thus, alternatively to an explicit model of an emotional state, a neural network is meant to set similarity conditions statistically. The process of establishing similarities, by constituting equivalence categories, is based on the massive repetition of the

---

<sup>1</sup> Play performed on November 21st, 22nd, and 23rd at the Center for the Arts at Enghien-les-Bains (95). Director: Christine Zeppenfeld; comedians : Valérie Le Louédec and Magali Bruneau; multimedia design : Alain Bonardi and Nathalie Dazin; music : Stéphane Grémaud; lights : Thierry Fratissier. Website : <http://www.latraverseedelanuit.net>

experience of the phenomenon. When the comedian is rehearsing or acting, the neural network "makes every effort" to recognize immediately the states of the comedian's voice. The quest for similarity continues.

In the two cases, learning and recognition, there is not any real dialog man-machine in the traditional sense of question and answer, but a frantic solicitation of one by the other. While learning, the neural network constantly appeals to the comedian; during recognition, the comedian is the one who appeals to the network relentlessly, to adapt her acting to what is stable and what changes. In both cases, one does not listen to the other, in the classical meaning of a dialog, but one works by continuously taking in information given by the other one.

#### **4.2 Instantiating a collaborative graphic generator**

The main goal of the device in this play is to make the comedian's voice control the character's "mental images," which are projected on the screen at the back of the stage, on a vast cyclorama (30 by 16 feet). The demands of artistic expression thus lead the designers to lessen the importance of instantiation, to give up the classical way of specifying a graphical problem.

Therefore autonomous agents were used to build the picture at the back of the stage in a collaborative way. These agents should be pictured as billposters who would work together to create a poster from fragments of images along with the sponsor's goals. The purpose is to create an artistic content by solving iteratively an optimization problem.

Each of these billposters is modeled according to a few variables, inspired by psychology, which correspond to different states of the voice (recognized by the neural network) which increase or decrease its "mood".

The results produced by this generative device that is given goals are amazing in terms of the distortion and movement of the image. Across performances, the movements on the screen are completely different, but they are always somehow "harmonious". This technical invention has also changed the relationship between the computer and the members of the project.

Some drama categories are thus displaced or questioned by this way of accessing artistic expression according to similarity and repetition. For a long time, the text of the play was long given as the only input towards expression and emotion [2].

Doing away with the text as the only medium, and especially by using the medium of digitalized voice, the processes that we have described, based on the establishment of similarities, pave the way for the creation of new drama emotions, and maybe new ways of theater creation and distribution with digital devices, provoking new uses for it. As with the distribution of musical contents by the Web, the auditory signal, its descriptors, the organization of "drama" contents with databases and MPEG norms that could be imagined, would reorganize the connection between theater and the audience<sup>2</sup>.

---

<sup>2</sup> We have partially studied this kind of reorganization in the case of digital opera, see [4].

## Conclusion

We have shown how the distribution of digitalized music does away with the dominant paradigm based on buying CDs for a dynamic indexation centered on the notion of similarity. Thus appears similarity-based calculus, acting on considerable masses of samples, which continuously change the balance of the man-machine dialogue, in an appeal started again and again to compare information. We have also shown how some artistic creations, heavily based on computers, proceed in the same way.

These prospects open many research avenues, building bridges especially between the search for content and artistic creativity. Great disruptions in the domains of culture and art, now intertwined by their massive digitalization, can be foreseen.

## References

- [1] Aristote, *Métaphysique*, Garnier Flammarion, 1993
- [2] Beurivage, Th., *Exploiter des données morphosyntaxiques pour l'étude statistique des genres – application au roman policier*, "Traitement automatique des langues" Journal, vol. 42, 2, text available at [www.atala.org/tal/](http://www.atala.org/tal/), 2001
- [3] Bonardi, A., Rousseaux, F., *Music-ripping : des pratiques qui provoquent la musicologie*, Musicae Scientiae, 2004
- [4] Bonardi, A., Rousseaux, F. *Composing an Interactive Virtual Opera : The Virtualis Project*, Leonardo Journal, Journal of the International Society for the Arts, Sciences and Technology, vol. 35, 3, p. 315-318, 2002
- [5] Donin, N., *Towards Organised Listening: Some Aspects of the Signed Listening Project at Ircam*, Organised Sound, Cambridge University Press, forthcoming in 2004
- [6] Drouillon, F., *Le calcul de la beauté : le principe de Pinocchio*, proceedings of the STP seminar, MSH Paris, 2003
- [7] de Gaulle, G., *La traversée de la nuit*, Éditions du Seuil, 1998
- [8] Pachet, F., *A Taxonomy of Musical Genres*, RIAO, Paris, avril 2000
- [9] Pachet, F., *Nom de fichiers : Le nom*, proceedings of the STP seminar, MSH Paris, 2003
- [10] Perrot, J.-F., *Des objets aux connaissances*, Journée Méthodes objets et Intelligence Artificielle : Frontières, Ponts et Synergies, Paris RIA, juin 1994
- [11] Rastier, F., *Sémantique pour l'analyse*, Masson, Paris, 1994
- [12] Rousseaux, F., *Pourquoi le Web resterait-il sourd à nos envies de musique et de son ?*, MEDIANET, Monastir, 2002
- [13] Rousseaux, F., *La notion d'œuvre à l'heure de la musique en ligne : comment concevoir la nécessaire médiation des affinités électives ?*, Klincksieck Editions, 2002
- [14] Simondon, G., *Du mode d'existence des objets techniques*, Aubier, 1989
- [15] Seridi, H., Akdag, H., *Approximate Reasoning for Processing Uncertainty*, Journal of Advanced Computational Intelligence, vol. 5, 2, Fuji Technology Press, 2001
- [16] Szendy, P., *Ecoute, une histoire de nos oreilles*, Éditions de Minuit, 2001
- [17] Vinet, H., Herrera, P., Pachet, F., *The Cuidado Project*, proceedings of the ISMIR Conference, Paris, 2002