

Assistance, advice and guidance with digital coaching

Jean-Claude Tarby

José Rouillard

Laboratoire Trigone, Institut CUEEP

Université Lille 1

59655 Villeneuve d'Ascq Cedex, France

tel. : (33)3 20 43 32 70 fax : (33)3 20 43 32 01

jean-claude.tarby@univ-lille1.fr

jose.rouillard@univ-lille1.fr

Abstract

We present in this article the concept of digital coach. More than a traditional assistance, the digital coach allows the user to improve his/her capabilities, by guiding him/her, step by step, particularly in a vocal way, and by stimulating him/her thanks to personalized information. Those are prepared specifically according to each user (texts and synthesis voices in his/her language) and are exploited by an application running on a mobile device. That makes it possible "to be coached", digitally, and individually, during the activity, and not preliminarily and in a generic way. The application fields are diverse (sport, cooking, do-it-yourself, culture, tourism...). A definition and a first taxonomy of digital coaches are proposed in this article. The concept is then illustrated thanks to a prototype of a coach embedded in a mobile device.

Keywords: Digital coach, personalization, assistance, human performance, voice synthesis, embedded application

1 Introduction

It is difficult today to have with oneself an application which can advise us, guide us, assist us, motivate us, etc., in our activities.. These acts are the prerogative of the coaches (Louart 2002) (Dufau et al. 2005) to whom the recourse has been in very strong progression for a few years, as well in the professional environment (management of team for example) as on a purely individual basis (management of the self-confidence, advise for relooking, etc.). These coaches adapt their expertise to accompany their clients in a personalized way. However, personalization is increasingly present today in computing: personalization by "skin" and by cascade style sheets (CSS), but also personalization of information by reserved access (login/password), by subscription (choice of topics on information websites), by "syndication" with RSS flows (Really Simple Syndication), by espionage (cookie, history of visits), etc. Mobile devices are also concerned by personalization: logo, ringing, screensaver, colours... are customizable.

These ten last years computing knew considerable progress in three fields which interest us more particularly in this article. The first is the mobility thanks to computing becomes pervasive and ubiquitous. The second is the assistance to the user, specifically

“intelligent assistances” based on “intelligent” agents, for example the Microsoft© “paper clip”. The third is the voice synthesis and speech recognition which can now be embedded on mobile devices. The evolution of the online help and “intelligent assistances”, the increasingly strong presence of personalization in interactive applications, the omnipresence of the mobile devices, and the continuously increasing quality of the voice syntheses, led us to propose in this paper the concept of *digital coach*. Our conviction is that the personalization techniques coupled with the intelligent assistance techniques should make it possible to carry out the *tomorrow’s coaching applications*. A digital coach is more than an intelligent assistance. Its characteristics are similar to those of a real coach, and it profits in addition to the advantages of technology. First this article presents the concept of digital coach. Then we present the prototype that we developed to show the feasibility of our ideas. Then, we conclude and give our prospects.

2 The concept of personal digital coach

2.1 Example: a coach of “jogging”

Here is an extract of a scenario to illustrate the concept of digital coach. Marc, a French man, decided to start jog again. He connects on a website for jogging with digital coaching. While creating his account, Marc gives personal data (sex, age, weight, height...), but also data on his former practice of sport, data on his quality of life (tobacco, drink...), etc. Then he indicates the training type that he wishes: desired time of training, level to reach at the end, etc. The website also asks him information about the place where he lives (a training in plain or mountain will not have the same contents!). Other information could be requested, for example if he prefers a male or a female voice (samples may be listened) and which kinds of music release or stimulate him. When this phase of information is finished, the system generates a training program. This program is composed of a detailed planning, as well as a set of MP3 files associated to each training day, that Marc will download from the website into his MP3 walkman. When the hour of training approaches, Marc goes to the place which was indicated in his planning, puts the ear-phones of his MP3 walkman, and listens to the MP3 files of the training. Here is an extract of what Marc will hear the first day, pronounced by the voice that he chose: “Hello Marc. Your goal today is to recover a sufficient rate of heartbeat to pass at the stages following days. We will start with a warming-up. Walk while inspiring deeply during 100 meters, then gradually accelerate the walk during approximately 1 kilometre, that is to say 15 minutes. To help you, you will hear music during the necessary time. This music will contribute to get you into step, and will accelerate to bring you at the desired speed”. Depending on the type of coaching that Marc chose, he may connect to the coaching web site after each jogging session to give information allowing the digital coach to better adapt its advises.

2.2 Definition of the digital coach

“The coaching [...] is defined as a continuous relation which makes it possible to the client to obtain concrete and measurable results in her/his professional and personal life. Through the process of coaching, the client increases her/his knowledge, improves her/his performances. Thanks to the interaction with the coach, the client clarifies

her/his objectives and get involved in the action. The coach's support makes it possible to the client to progress more quickly towards the desired results because the relation of coaching invites her/him to be centred on her/his priorities and to be conscious of her/his choices"ⁱ. Our work consists in implementing a "real" coach in the form of a digital coach in a computing system; this system can be an application, a website, or any other type of computing system. The coaching that we implement in our work is the exact translation of the previously given definition. The coach and its client (i.e. the user who wants to be coached) "know themselves" and "meet". According to the requested type of coaching (embedded, intrusive, interactive...) the user can have more or less regular contacts with her/his coach in order to refine her/his coaching. This can be done explicitly (at the request of the user) or implicitly (the actions of the user are returned automatically to the coach who then makes decisions as for possible modifications for the coaching). Like a "real" coach, the digital coach provides personalised advises to its clients. So, from the same request, the digital coaching will be different for two persons.

The digital coaching such as we define it in our work can be applied to various fields like sport, cooking, do-it-yourself, culture, tourism, etc. We include in our definition of the digital coaching the previously defined concept of coaching but also the concepts of guide, assistant, advising, supervisor, and "shift-person" (see §2.4 for the example of the travel). In our work, the use of voice is very important. As explained in §2.4, voice allows to produce personalised advises which should be perceived by the client better than textual information.

2.3 Characteristics of the digital coach

Some characteristics can be defined for digital coaching:

Expert of the field. Like a real coach, the digital coach must perfectly know its field. This characteristic is not specific to the digital coach. More, it is external to this domain. All the techniques of knowledge management can be used to work out this expertise: expert systems, neuronal networks, bayesian networks,...

Personal relation with its client. The digital coach and its client are linked by a privileged personal relation. Thanks to this relation the coach to adapt its speech as on the form (turn of phrase, intonation, style of music...) as on the substance while adapting to the real context of the client (named place for a sporting exercise, client's material for do-it-yourself, etc.) but also to her/his habits and preferences, etc.

Embedded/centralized. A centralized coach is on a web server or any other type of centralized application. On the other hand, an embedded coach is a coach that one easily carries with oneself thanks to mobile devices (MP3 walkman, PDA, mobile phone, etc.).

Reactive. Since it is possible to store the history of the coaching and consequently to know its evolution (client's progress and regressions, rate of success, etc.), the coach can still better adapt its speech and its actions with respect to the user.

Static/dynamic. A static coach generates its recommendations before they are used, for example in the case of the jogging in §2.1. On the contrary, a dynamic coach chooses,

ⁱ Translated from International Coach Federation France, <http://www.icffrance.org>

adapts and/or generates them in real time, possibly while basing on dynamic data such as the GPS position of the user, the number of people who are present with the client in a room of museum, etc.

Intrusive or not. According to the wish of the user, the coach can be intrusive or not. In the first case, the coach can intervene without user solicitation, whereas in the second case the user must explicitly call the coach.

Interactive. The interactivity with the coach is of two types. On the one hand, in certain situations, it is interesting that the coach proposes choices. In this case, the user will choose according to her/his desiderata or the context of the activity. For example, for a coaching on the assembly of a piece of furniture in kit form, the coach can say: “If you are alone, listen to the next track. If you are helped by somebody, listen to track 5”. In addition, the user will be able to command her/his coach with voice (including with a PDAⁱⁱ) with orders such as: “read all” (automatic reading without touching the keyboard... practical when the user has the hands occupied), “read stage 5” (access to a stage by its direct number), “read the stage speaking about...” (access to one or more sequences containing a keyword), “stop/pause/read”, etc. These vocal orders can be supplemented by a graphic user interface.

2.4 Advantages of the digital coaching

The majority use of the voice in the digital coaching (by voice synthesis or from real recordings) is of primary importance. Hearing instructions rather than reading them releases the sight and possibly the hands which allows to use them for other tasks (interesting for the sport, do-it-yourself, or cooking...). Moreover, the voice synthesis made huge progress, and today the synthetic voices are able to produce emotions almost as well as human voices^{iii, iv}. Lastly, hearing a “human” voice talking personally to you should produce more emotions than a text with impersonal instructions.

We created a system containing voice synthesis which is able to talk in several languages. Our digital coaching is thus *multilingual*^v and *polyglot*^{vi}. It is *multilingual* because we have synthetic voices for various foreign languages, allowing pronouncing a text as a native of the country would do it. Our system is *polyglot* because it can pronounce sentences mixing various languages. These two characteristics are interesting for example if the user must travel in a country of which she/he does not master the language. A digital coach can help this person to prepare her/his voyage, for example by giving her/him information about the administrative formalities to achieve, but a digital coach can make much more by helping the user to communicate at critical phases during the trip, for example to take a taxi or to obtain the room that she/he booked. For instance, if the digital coach is combined with the travel company that the client used, it can know where the client will arrive (i.e. the name of the airport), which hotel (and details about the room) he booked, and so on. In such a situation, the digital coach can

ⁱⁱ <http://www.microsoft.com/windowsmobile/downloads/voicecommand/default.mspx>

ⁱⁱⁱ Loquendo, http://vhost.oddcast.com/vhost_minisite/support/ssml/expressive_cues.php?voice=4

^{iv} Miriam et Léon, <http://www.timespace.com/french/miriam.asp>

^v Multilingual: someone who speaks three languages or more, learned as mother tongues.

^{vi} Polyglot: who speaks several languages.

adopt various strategies. First of all, the digital coach can be used as “shift-person” while speaking instead of the user. For example while entering the taxi, the user runs her/his coach which will say to the driver “Take me to the Hilton hotel, 3000 Paradise Road” with a perfect pronunciation. This is possible because the digital coach prepared some sentences with regards to the client’s travel. In the same way, at the hotel the user will get his coach to say: “Hi! My name is Marc Denis. I've reserved a room. The confirmation number is XNK2078”, etc. Again, this is possible because the digital coach received this information from the travel agency.

The second strategy consists for the coach to encourage the user to speak. For example the coach can propose training over several days before the departure. This program just like the preceding example about jogging will be based on a thorough knowledge of the starting level of the user and the level to reach.

With the last third strategy, the coach can accompany the user at the convenient period. If the user followed the training evoked above, or if she/he estimates that her/his level is insufficient, the coach can play the role of prompter at the opportune time. Thus, when the user must pronounce a key sentence (for example “Take me to... ”), she/he will ask her/his coach to whisper the sentence to her/him, this in order to have a non-disturbing sound reference mark.

3 Current realisation

In order to validate our ideas, we produced a generic prototype, which is not dedicated to a specific activity. For the moment, this prototype corresponds to the vocal part of the coach (production and use). We can use it in various fields such as sport, cooking, car mechanical, etc. Therefore, it must be connected to an application that is expert of the associated field.

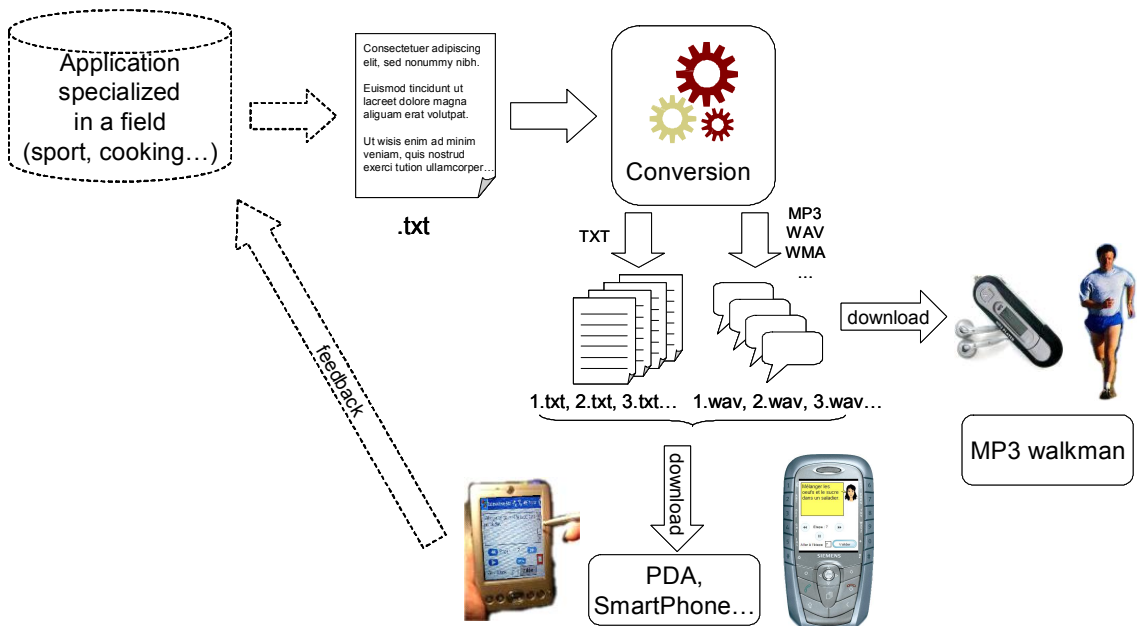


Figure 1 : General principle of the digital coach

Our prototype consists of two applications. A first application on PC generates the necessary files for the task to be realised, and another, embedded on a mobile device, uses them. The prototype was developed in C# with the *ActiveTTS*^{vii} library and the library used for the voice restitution is *Hekkus*^{viii} for Pocket PC. The prototype process occurs in two times (cf. figure 1). First, one provides a textual file to the PC application, including as many paragraphs as stages to realise. Then the application generates N textual files (.txt format) and N vocal files (.wav format for example). The user can select the design features to apply to the voice synthesis. In our prototype, we can specify the kind of voice (man, woman, child...), the language (French, English, Italian...), volume, speed, and the desired audio format (8 Khz, 8 bits, mono, for example). In the long term, these choices will be carried out in a (half)-automatic way, according to the task and certain contextual elements. The produced textual and sound files are then transferred on the mobile device. In our study, we used a HP iPAQ HW6500 Smartphone. The second application is embedded in the mobile device and exploits these data by displaying for each stage the associated text, and by diffusing the voice synthesis which accompanies it. This software is more than a simple sound files player. It simultaneously presents elements of a same stage and makes it possible to stop and (re)play a voice synthesis, jump to another file representing another stage, change the language of the interface, etc. It will be easy, in a forthcoming version to add an illustrative picture for each stage (.gif animated format for example).

4 Related works

The “intelligent assistances”, based on task/user models, which aim to help the users in the realization of their tasks, are very close to the concept of coaching. Even if many work were devoted to this approach (Jameson 2003) (Fink et al. 1997) (Langley 1999), the results are still not very convincing as for the taking into account of the individual and very varied criteria users (Capobianco et al. 2006). Approaches in connection with the artificial intelligence such as Bayesian networks (Nicholson et al. 1984), neural networks, or multi-agents systems (Wooldridge et al. 1995) (Kolski et al. 2004) (Eisenstein et al. 2002) appear to be more effective. It seems obvious that the digital coaches’ expertise must take this kind of systems as a starting point (Gonzales et al. 1998). However, these systems are complex to implement. They can be carried out only on desktop computers (not on small mobile devices), and they are not dedicated to a particular person (no personalization, no management of conversations history...). However, more and more products are developed around the coaching concept^{ix}, but none goes until a mobile personal accompaniment as described in this article. On the communication level, some systems like Collagen (Eisenstein et al. 2002) use text-to-speech synthesis and voice commands, and are equipped with capacities of animation (character movement, facial expressions, pointing elements on the screen, etc.). Approaches such as those of the Animated Conversational Agents (ACA^x) (Cassell et

^{vii} <http://www.guangmingsoft.net/activetts>

^{viii} Hekkus project, <http://www.shlzero.com>

^{ix} For example, Polar Running Coach, www.polarrunningcoach.com.

^x GT ACA, Groupe de Travail sur les Agents Conversationnels Animés, <http://www.limsi.fr/aca>

al. 2000) or of communication languages such as KQML^{xi} (Finin T et al. 1994) are tracks to study seriously in order to answer questions such as how to make agents more human or/and more communicating with humans. Concerning the vocal issue, in industrial or research projects, it is mainly used in output (voice synthesis) (Carbonell 2006) (Kieffer et al. 2006), sometimes in input (vocal commands), and seldom in a multimodal way, like originally exposed with the « Put that here » concept (Bolt 1980). Commercial products are developed more and more around vocal for mobile devices, for example, cities tours^{xii} or electronic checklist systems^{xiii}, but they are only playing sound files and do not integrate the coaching dimension such as defined in this paper. Concerning the adaptation to the activity context, some systems are able to adapt themselves to modifications produced by the user (change of road, change of heartbeat...). It is the case, for example, with car GPS or within the MPTrain system (Oliver et al. 2006). Decisions are made by these systems and are given to the user in an oral, musical, graphic or textual mode. But these systems are not coaches such as we defined it here. They give contextualized but not personalised information.

5 Conclusion and further works

We proposed in this article the concept of digital coach. We showed that digital coach, and digital vocal coach more particularly, makes it possible to have a personal assistance which one can carry with oneself on mobile supports. This assistance includes capacities of assistance, motivation, guide, advices, etc. We drew up a first list of properties characterizing the digital coaches, as well as the list of the advantages of this kind of coaching. We gave examples of applications and explained the way our current prototype works. It shows practicability on the vocal level. For the prototype, we are studying the establishment of the vocal commands, as well as the textual and vocal indexing in real time. In the long term, one can imagine that the digital coach will be connected to sensors (pulse, light, heat, etc.) to adapt itself in real time to the context and the user's reactions (Oliver et al. 2006). Works on intelligent agents (Kolski et al. 2004) and on man-machine dialogue (Sadek 2004) should also make it possible, in the near future, to overcome a little more the border which separates digital coaches from real coaches.

6 Acknowledgments

This work has been partially financed by European FEDER funds and is a part of the MIAOU project (Région Nord Pas de Calais).

7 References

Bolt R A (1980) Put that here. Proceeding of SIGGRAPH'80, ACM Press, pp. 358-364.

^{xi} KQML, <http://www.cs.umbc.edu/kqml>

^{xii} www.pocketvox.com

^{xiii} www.avionix.com/checklist.html

Capobianco A, Carbonell N (2006) Aide en ligne à l'utilisation de logiciels grand public : problèmes spécifiques de conception et solutions potentielles. In Boullier D et al. (eds). Systèmes d'aide opératoire. Octarès, France.

Carbonell N. (2006) Ambient Multimodality: towards Advancing Computer Accessibility and Assisted Living. International Journal on Universal Access in the Information Society (UAIS), DOI: 10.1007/s10209-006-0027-y.

Cassell J, Sullivan J, Churchill E (2000) Embodied Conversational Agents, MIT Press.

Dufau H, Perdriset F (2005) Le coaching, enjeux, paradoxes et perspectives. La revue Communication et Organisation, n°28, 2005.

Eisenstein J, Rich C (2002) Agents and GUIs from Task Models, In Proceedings of Intelligent User Interfaces 2002, San Francisco, CA.

Finin T et al. (1994) KQML as an Agent Communication Language, 3rd International Conference on Information and Knowledge Management (CIKM94), ACM Press.

Fink J, Kobsa A, Nill A (1997) Adaptable and adaptive information access for all users including the disabled and the elderly, In Jameson A, Paris C, Tasso C (eds.), Proceedings of the Sixth International Conference on User Modeling, UM97, Chia Laguna (Italy), Wien-New York: Springer, pp. 171-173.

Gonzales C, Wuillemin P.H (1998) Réseaux Bayesiens en modélisation d'utilisateurs, Sciences et Techniques éducatives, Vol.5 n° 2, pp. 173-198.

Jameson A (2003) Adaptive Interfaces and Agents, In Jacko et al. (eds.), Human-Computer Interaction Handbook, Mahwah, Lawrence Erlbaum Inc., pp. 305-330.

Kieffer S, Carbonell N (2006) Oral Messages Improve Visual Search. In Proceedings of ACM International Conference on Advanced Visual Interfaces, Venice, Italy.

Kolski C, Grislin-Le Strugeon E, Adam E, Mathieu P (2004) Conception des systèmes multi-agents : pistes de réflexion en vue de futures coopérations entre ergonomes et informaticiens, In Actes du congrès ERGO-IA'2004, ESTIA, Biarritz, pp. 147-154.

Langley P (1999) User modelling in adaptive interfaces, In Proceedings of the Seventh International Conference on User Modeling, UM99, Banff (Canada), Wien-New York: Springer, pp. 357-370.

Louart P (2002) Le coaching : son intérêt, ses méthodes. Les Cahiers de la Recherche. Claree, IAE de Lille, 2002.

Nicholson A, Korb K (1984) Bayesian Artificial Intelligence, Addison-Wesley.

Oliver N, Flores-Mangas F (2006) MPTrain: A Mobile, Music and Physiology-based Personal Trainer, Microsoft Research Technical Report MSR-TR-2006-31, March 2006.

Sadek D (2004) Nouvelles perspectives pour l'ergonomie des interactions personne-machine : dialogue naturel et agents intelligents, In Actes du congrès ERGO-IA'2004, ESTIA, Biarritz.

Wooldridge MJ, Jennings NR (1995) Intelligent agents: Theory and practice, Knowledge Engineering Review 10(2).