Nico Baken, Edgar van Boven, Ramin Hekmat and Ludmila Menert

Virtual Mobility Enabling Multidimensional Life

Each day, one day is added to our past, and thus one day subtracted from our future. That is why the future ought to be simpler than the past. Why is it then, that we always wrestle with the future? Because the new and changing concepts are increasingly complex.

In this paper, we endeavour to describe some novel and intricate concepts in the ICT world such as presence, virtual mobility and avatars. They will have a crucial impact in our economic and social lives. These themes are positioned and explained as a logical evolution, a concatenation of breakthroughs in man's capabilities, such as speech, counting, writing, and the first steps in (tele)communication. The evolutionary sophistication in (tele)communication can be caught in three main parameters: distance, time and richness. These three entities behave in time as communicating vessels, but the 'ceiling' of the product of the parameter-values increases steadily. Furthermore, an identity matrix is described, highlighting an explosion in telecommunication far beyond human beings.

Historic Overview and Future Image

At the end of the last Ice Age, man, who still lived in small tribes of hunters and collectors, started to establish village communities of farmers and shepherds. They became colonists. With the settling, the registration and reproduction of numbers – say quantities of grain, olive trees, sheep, etc – became a necessity. In short, we had to learn how to count: to symbolise numbers. The need to map these numbers on one or more persons led ultimately to the arithmetic process of division.

The reproduction of symbols for numbers has led to the alphabet. Thus, counting was the basis for writing and knowledge transfer across the barrier of time itself! In writing, in fact, man developed his first primitive time machine. However, the richness of this communication is low, it is asymmetric, text based and therefore poor in 'senses' (compared to the richness of face-to-face communication, where much more information is carried in visual and acoustical cues); nevertheless a skill of tremendous importance contributing in an acceleration of the knowledge build up and progress of mankind. After all: indeed script is what enables us to register our knowledge and part of our experiences and pass them on to next generations. We pay a price for this: the gift of oral transfer over many generations is diminishing. With each innovation, something seems to get lost, and in that sense progress is relative. The power of the time machine of script (and music scores) was exponentially magnified by the arrival of the art of printing in 1450: the basis for decentralisation of knowledge. Here, too, a trade-off took place: knowledge versus insight and memory. In the field of human communication, we see the following major innovations: language as a step from non-verbal to verbal communication, a miracle that took more than a million years; counting and script, THE innovations of the last ten thousand years. But were they sufficient?

No, the need evolved for real-time communication over longer distances, as people themselves became more mobile. The physical distances created a barrier which could not be conquered quickly enough on foot, on horseback, and later by car, train and airplane. Around 1864 Alexander Graham Bell envisaged that his famous patent would be merely used to listen to music from a distance. Apparently, he did not relate his patent with the possibility to realise a commercial public telephony service, because he simply could not imagine that people would be willing to talk to one another without their physical presence and even pay for it.

Bell probably underestimated that for human speech, real-time bridging distance could be more important than the sacrifice of the familiar input for the human eye. Now we discover the pattern. As with *communicating vessels*, we gain in one parameter but must concede in another one: distance is conquered but we lose in richness. Richness comprises sub-parameters such as symmetry and senses; in the case of telephony, the price is paid by missing the physical presence and all involved senses; that is, sight, smell, etc.

In the 20th century, developments accelerated: a telephony network branched out through the world, in which technical

Authors
Nico Baken
Delft University of Technology and KPN
Oosteinde 187
2271 EE Voorburg
The Netherlands
Tel: +31 70 3876378
Email: n.h.g.baken@kpn.com

Edgar van Boven
KPN
Email: e.f.m.vanboven@kpn.com

Ramin Hekmat
Delft University of Technology
Email: r.hekmat@ewi.tudelft.nl

Ludmila Menert
Consultant
Email: ludmila.menert@fluency.nl

innovations came in rapid succession. We moved from manual switchboards to electromechanical ones and ultimately to computer-operated exchanges with an exponentially increasing number of users†. The fixed and mobile telephony network now totals roughly three billion users. Even the most rational of us must experience some emotion at this: an unprecedented accomplishment, because it works and works (nearly) all the time! Applications for new connections, repairs, billing, etc: it has all been taken care of! But is it enough?

No! Despite the arrival of mobile telephony, which has progressed much faster than ever expected in the past 10 years, and despite the unprecedented popularity of the Internet, it is NOT enough. Distance and time barriers have been, as explained, partially overcome. The transfer of sound, text and images in a limited quantity per time unit – that is, narrowband – is possible. High-quality real-time video and large data files require speeds higher than 2 Mbit/s, and this type of broadband information cannot yet be transferred at anytime, from and to anywhere end to end, customised to suit the needs. This must be the next step, and once again it is a giant one because it cannot be made using the existing infrastructure. It will lift the ceiling of the product of the parameter values: time, distance and richness. Until now, we have experienced the following sequence of giant steps in the development of communication:

Non-verbal communication	→	Language
Language	→	Counting
Counting	→	Script
Script	→	Narrowband telecommunication

The predicted subsequent steps will be:

Narrowband telecommunication	→	Broadband telecommunication*
Broadband telecommunication	→	Virtual reality and virtual mobility

The innovative steps to broadband telecommunication and virtual reality and virtual mobility appear to be predicted and scheduled for the next 15–20 years. Each step is necessary for the next one and is prepared for by smaller intermediate steps. Thus we discover that natural barriers such as time and distance are partially overcome through script and the first steps in telecommunication. In this process, concessions are made in the 'richness' of the communication session. With increasing bandwidths and computing power this can be relaxed and/or repaired 'virtually'. In this sense, the pitfall of Bell can repeat itself for us. Some of us may not be able to comprehend the presence of the communicating party or parties without them physically being there. And, for the factor time (difference), it is hard to grasp that travelling through time does not actually take place, but communication with deceased living entities or no longer existing non-living entities, will take place by means of systems that represent them. Obviously, the sacrifice in richness is overcome by yet crossing another barrier. The sacrifice is softened or repaired when a major innovation takes place and the ceiling is lifted again.

At the dawn of the 21st century a well-known future image loses its science fiction character: You can appear anywhere in a digitised representation of yourself or what you want to be associated with. You and your avatars (described in the next section) can have real-time communication sessions without physically transporting your body, when using the right ICT means. We believe a substantial part of mankind will adopt this concept of virtual mobility during the first half of this century. In general, a lot of people wish to enjoy a more efficient and richer multidimensional life and be more successful than others. The considered dimensions in 21st century life are:

- the good old physical life;
- life in cyberspace being yourself;
- life in cyberspace being your enhanced self, or even faking to be somebody else; and
- life in cyberspace 'meeting' other non-human (id)entities

Figure 1 shows an image of physical life in 2014. The yellow beams symbolise information flowing in cyberspace. In Figure 1 people and applications (residing on a plethora of devices) continuously interact. (Nearly) all services are enjoyed wirelessly. Radio ranges will be shorter due to higher bandwidth.

Really new compared to today's situation is the multimedia streetlight on the curb. It not only offers society wireless connectivity for real-time communication, but provides security services, personalised advertising and location-based information services as well. The deployment of intelligent sensors and ICT (id)entities will fundamentally change society.

Figure 1 shows a lady in a wheelchair triggering a sensor that adapts the stairway. A blind man is guided by a multimedia cane, telling him where to go. His health parameters are continuously monitored and sent to his medical account. His doctor asks for an overview of the blood pressure of all his patients on his flat screen. The dog on a leash has an ICT identity, too. It can always be traced when it's lost. The dog's health can optionally be monitored as well.

The man sitting at the sidewalk cafe on the right is playing a multimedia game with the girl sitting on her balcony (on the upper right). The waiter establishes a paying transaction. No banknotes are involved anymore. The car is equipped with an

Figure 1 Future image[7]

† The number of telephony connections throughout the world has increased from the earliest times by 6.3% per year. Expressed as a formula: $A_n = A_0\,(1.063)^n = A_0 e^{0.061n}$. A power of e with an extremely small exponent. The doubling period N in years = $\ln(2)/\ln(1+x/100)$, in which x is the growth percentage per year, renders N = 11.3 years. Our thanks to J. W. Meijer.

* In terms of services, this is the transition from narrowband telephony and best effort data to carrier grade multimedia.

online vehicle area network that can temporarily host every personal area network of all the people who take a ride, not only the owner of the car.

Realising the future image depicted above requires an ICT identity architecture in order to discriminate and service all continuously interacting entities. Our (id)entity matrix, Figure 2, tries to grasp all thinkable future information flows in cyberspace, by means of typical examples of interactions between (id)entities.

An important development is the rise of (wireless) communication devices (D). Many different types of devices will be able to (autonomously) communicate. They will outnumber people before 2010 and generate and process more information than people. In the (id)entity matrix in Figure 2 devices (D) are discriminated from all other non-living entities (E) because of their communication function. All (id)entities (A–F) depicted in Figure 2 are expected to be equipped with:

- sensors being their (additional) eyes and ears; and
- applications and actuators being their (additional) brains, mouths and limbs (for example, cat door (E) opens when an authorised cat (A) approaches).

Passive devices (like sensors and electronic tags) will outnumber active devices (D).

Two different long-term images exist:

1 *Autonomous devices gaining importance,* being smarter, faster and far more efficient then average people for specific tasks. In 5 years time, devices and (their) applications will be able to take simple decisions; in 10 years time they will take more complex decisions (helping their owner) without asking direct permission. Being connected to vast information databases, devices/applications can be consulted real time if you can afford it.
2 *People remaining in the lead,* physically enhancing themselves with technological extensions and brain-controlled user interfaces. People start to engineer themselves (in order to be more successful than others who don't use these features).

No matter which of these two images will dominate, the simplicity of user interfaces will be crucial.

(Id)entity matrix example:

In 10 years time, water (E) will be more expensive than today. In 2015, a group of apple trees (F) is thirsty. Due to hot and dry weather going on for days, the trees' humidity sensor (F) tries to trigger the irrigation system (E) requesting for some precious water. The communication application of the irrigation system (E) queries the weather report domain (B). Weather satellites forecast rain in five hours. The irrigation control system decides to wait. Somewhat later the trees finally enjoy the welcome rain, and the farmer (B) a lower water bill.

Figure 2 Information will flow among all (id)entities in this (id)entity matrix (panta rhei)

from \ to	A Animal	B Business	C Consumer	D Device	E Entities (non-living & non-D)	F Flora
A		Show their condition	The cat's location	This cow is hungry	Not applicable	Not applicable
B	Remote tag info	We could be partners 	Buy our Product !	Install software	Archive search	Guarding rare trees
C	Come and get it !	Your product doesn't work	 Bla bla bla	I'd like to download	What is your price?	Looking at a rare plant
D	Dinner Time !	Somebody stole me !	Your car needs oil	Which codec? 	Search query	Checking humidity
E	Door opens	Camera images	Camera images	Water temperature		Not applicable
F	Not applicable	Help, I'm cut down	I'm at the Hortus Botanicus	I could use some water	Not applicable	

Virtual Mobility and Avatars

Today's real-time presence information comprises the 'MSN pawns' and maybe a few other icons on communicator applications. Most instant messaging (IM) users today only share with their community of interest whether they are online or not, typing additional status information themselves.

It is technically feasible in the midterm to distribute real-time status information revealing the actual state of all used ICT means and in the long-term even physical events. Figure 3 shows a schematic presentation of this concept.

The main three messages of the presence concept shown in Figure 3 are:

- There will arise a gigantic new traffic flow caused by people continuously sending their real-time status information to their communities.

Figure 3 Next-generation presence concept

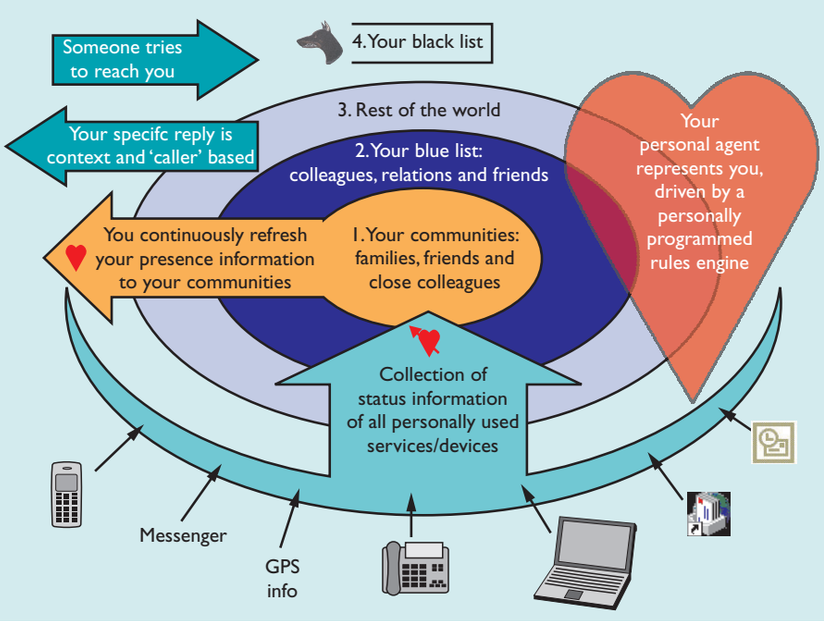
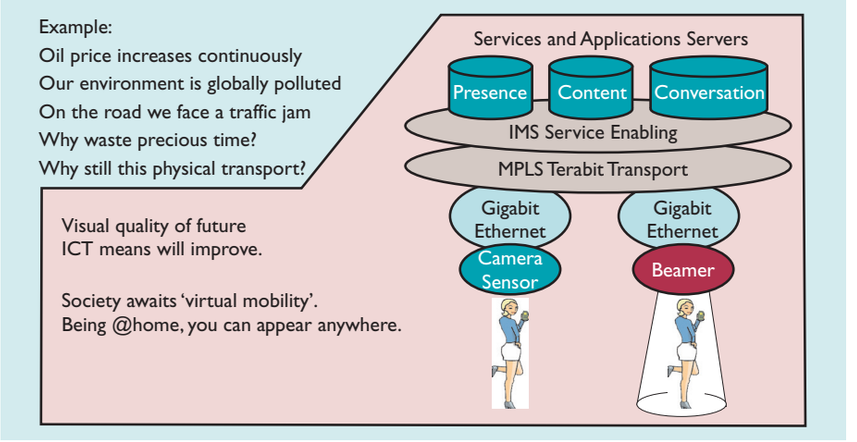


Figure 4 **Example stage 3 virtual mobility benefits**



- People will continuously define and redefine their communities (your tribes of virtual nomads)
- It will bring telcos and service providers a new revenue-generating feature in their packages.

Status information of all ICT means will/can be aggregated:

- your geographic location (Global Positioning System);
- being active in a telephony session (fixed and wireless telephony (video or sound only));
- being active in 'don't disturb mode' (for example, reading a document, writing towards a deadline);
- IM or email sessions (messenger); and
- Outlook (your schedule).

Future presence information will be spread to your communities in a context and preference based way, depending on:

- time of day;
- your current activity (working, enjoying holiday, sleeping, eating, shopping, etc.); and
- your current indicated mood, health and energy level (sometimes you will be in 'leave me alone mode').

When you adopt this next-generation presence technology you will have to accept the continuous monitoring and spreading of your status, and to learn to program your rules engine. Today people have got used to filling their schedule in Outlook. For market acceptance of this new presence concept, simplicity, privacy, and user friendliness of rules engines will be crucial.

Probably your communities only comprise a minority of all people/entities interacting with you. Depending on who or what is trying to reach you, your actual preferences, context and your status, they will receive your tailor-made reply. It could look like:

1 Members of your communities (continuously receiving your presence information) will be met as soon as possible and/or even directly and above all warmly and with your personal touch dedicated for this special person.
2 Your blue list will consist of precious relations, colleagues and friends to be contacted as soon as possible, friendly and with a personal touch.
3 All unknown 'callers' probably receive a standard, polite, friendly, open answer (comparable to your current voicemail greeting).
4 People on your blacklist receive a neutral voice or video mail. Don't call us we call you!

In general the service evolution could likely follow these stages:

1 Today's interactive multimedia and presence (mainly Microsoft messenger based).
2 Today's presence concept transforms into 'virtual mobility' in a parallel cyber world.
3 Virtual mobility concept will be enhanced with 3D projections (being at X you can appear at Y) (see, for example, Figure 4).

It means a paradigm shift: virtual mobile people and their representatives in cyberspace (avatars) will generate an enormous new traffic flow. It will consist of:

- traffic among avatars,
- traffic among people and avatars,
- availability information (presence information to your communities),
- two- and three-dimensional images and projections of people communicating real-time, and
- traffic caused by meeting non-human (id)entities (A, D, E and F).

The evolutionary path to ergonomically feasible avatars is not trivial. Below we elaborate on the technological implications of user-friendly avatar design. A personal assistant should help instead of irritate people.

Avatar types and designs

Avatars are used in virtual environments to act on behalf of *all* (id)entities (see (id)entity matrix Figure 2), as their virtual, mostly personified, representation. The avatar interface should offer comfort, convenience and efficiency, being unobtrusive and yet effective. This implies use of sensors in order to provide information about the entity itself (humidity/food/drink/health monitoring), but also to function as the ears and eyes of such devices in their interaction with humans in the physical world. So-called attentive appliances are able to assess whether or not to alert the owner (for example, to an incoming message) not only on the basis of the user's presence status, but also by monitoring the user's current activity, level of attention and even information load[2]. To be able to do this, the avatar's skills need to include reasoning: the value of a message is weighted against the cost of the disruption and the appropriate action is taken. Also, the appropriate channel should be chosen depending on the information type and the situation. Voice-only interfaces should be deployed either when voice access is the only possibility (car, hands/eyes needed for something else, etc.) or when talking is faster then typing and the content is fit for being listened to (long texts, complex instructions, overviews, lists or tables are unsuited for auditive presentation due to sequential processing and the limited short-term memory in case of auditive input).

A personal agent is designed and customised by the owner in order to reflect his or her actual or assumed persona. This enables you to improve your representation, by enhancing or adjusting your image. You can even choose to be represented by an image or persona that is totally dissimilar to yourself. As using such a *mask avatar* can bring the avatar owner psychological and emotional (though possibly addictive) satisfaction, we can expect growing demand for this type of fully customisable personal agents. However, the attraction of this aspect of avatars and cyber life in general is also a danger. Avatar adoption might cause tendencies towards withdrawal from physical society.

The appearance and behaviour of agents will always be designed with the main functions of the interface in mind: lively anthropomorphic agents may enhance the perceived presence, creating a favourable attitude towards advertisements, and the willingness to revisit a site. An anthropomorphic agent with some carefully crafted 'social' behaviour can have a positive effect on the user experience; for example, in interaction with patients. For specific tasks, users may prefer a machine to a human when the machine provides at least the same level of service and certainly when the machine can help quicker than a human (automated attendant, online banking), or better than a human (for example, Google).

Figure 5 Schematic representation of human–machine or machine–machine interactions

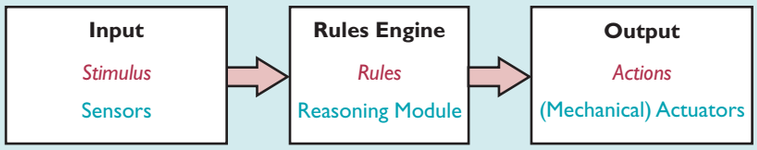
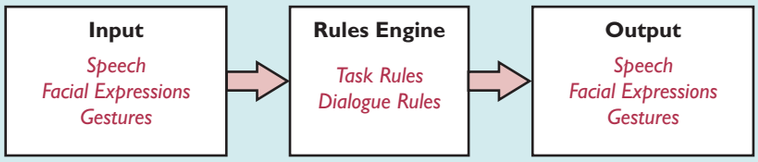


Figure 6 Schematic representation of human–avatar conversational interactions



Avatars powered by speech technology and artificial intelligence

As explained earlier, we expect even non-living entities from the identity matrix in Figure 2 will be interacting. For any interaction, the schematic representation shown in Figure 5 can be used.

In our earlier example of the thirsty trees, the stimulus (triggered by a particular value of the sensor indicator) is sent to the irrigation system, which then follows its rules (here carefully defined so as to take water shortage into account) and takes action by first sending a query to the weather satellite. (Simpler rules would just send a request to the irrigation system directly.)

When an avatar communicates with humans, it needs to emulate human communicative behaviour. This has some far-reaching technological consequences.

When we extend Figure 5 to describe such a conversational session, the input will be natural speech of a human and the output human-like speech of the avatar (Figure 6). However, speech is not all: in human face-to-face conversation, the verbal content of speech is supplemented by non-verbal information in the form of acoustical cues like stress, loudness, pitch, rhythm, and in the form of gestures and facial expressions, signalling topic-comment distribution (what are we talking about and what is being said about it), relevance, deictic reference (pointing) and attitude.

The most challenging module, though, is the *rules engine*: combinations of computational linguistics, artificial intelligence (AI) and knowledge management are needed for the application to do its work: first, to carry out its task (for example, provide information), and, second, to maintain a human-like dialogue with the user.

Figure 7 shows the different technologies that need to be combined for a conversational session between a human and a human-like avatar.

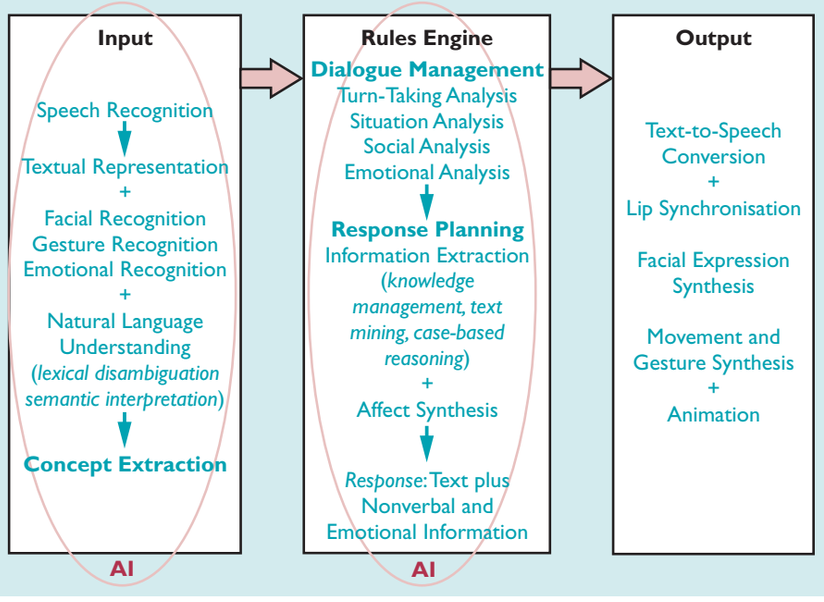
Correct input interpretation and response planning require context analysis on various levels: linguistic, situational, social and psychological (emotional). For example, correct recognition of the user utterance that sounds like 'seventy four' as 'seven to four' and its consecutive interpretation as '15:53 [time]' is facilitated by intelligent use of relevant contextual information. The same applies for interpretation on a more abstract level, for example substituting 'John' for 'he' in the sequence 'Reschedule the meeting with John. He can not come today', or, on yet another level, substituting 'Tom' for 'my son' in 'I want to call my son'. Information from previous sessions is needed to correctly interpret the emotional message expressed by above-average loudness and pitch values accompanying the phrase 'Not again!' The same levels can be distinguished in system output, where responses need to be generated in stylistically appropriate language, syntactically and semantically correct utterances (resolving pronunciation of words with identical spelling which are pronounced differently depending on their meaning) and naturally sounding speech.

Allowing for so-called 'free speech' input means not only that the speech recognition engine can interpret natural language utterances† (which in itself already implies quite advanced AI), but when the input can be anything, the engine has very little to base its decisions on to resolve ambiguities on different levels, as illustrated above.

Understanding the typically short, often incomplete and grammatically incorrect, utterances in free speech depends heavily on knowing the context of the conversation. The system needs the support of pragmatics. The approach taken in most current implementations is to use relatively simple pattern matching, based on domain knowledge. This artificial intelligence is still far from 'human': inserting words or whole parts of sentences from the user's comment in the responses, which are often randomly

† *Natural Language Processing (NLP)* uses a syntactic, semantic and pragmatic analysis of textual or speech input in order to allow for a free and flexible exchange in written or spoken natural language. Models for semantic equivalence and entailment are deployed for the correct interpretation of overtly distinct lexical content. Pragmatic knowledge, that is 'knowledge of the world' or knowledge of the domain of the discourse, is used to resolve ambiguities and make correct deductions from the available information.

Figure 7 Integration of different technologies required for human-like conversation



chosen from a set of arbitrary sentence templates, or responding with a predefined answer when a specific word in the question is recognised makes such a so-called *intelligent agent* just more or less an extensive pattern-matching case-statement. Complex discussions about complicated subjects are beyond the state of the art of AI.

Such systems can be improved semi-automatically by so-called *ripple down rules*, which provide a form of teacher-guided incremental learning method. Another possibility is so-called *batch learning* by induction over examples: logs of user interactions are stored and collected and machine learning programs seek common patterns and construct sets of rules (or decision trees) that describe the preferred behaviour of the system.

A further step will be when avatars themselves will be able to access knowledge stored in different locations on the Internet. Information about the user (past interactions, preferences, etc) could be delivered to the avatars by the personal assistant of the user. Such integration would help to create a self-learning intelligent environment where information gathered about the user will become available to other users†.

Information flows

This means information exchange and information flows between avatars: in addition to communication with people, avatars will increasingly communicate with each other to improve themselves. In the future, the self-learning ability of avatars will outgrow the probabilistic machine learning; avatars will become more autonomous and able to take more complex decisions and eventually organise themselves in hierarchical layers.

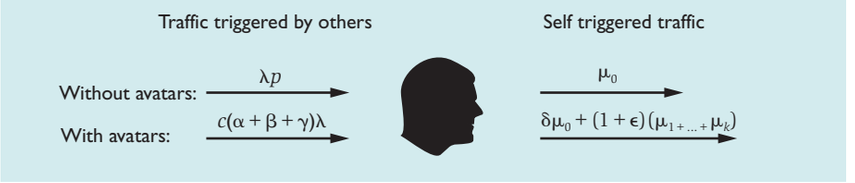
In effect, human–machine and machine–machine communication will dwarf the traffic volume of inter-human communication.

Proof of traffic increase caused by virtual mobility

Figure 8 shows a simple scheme proving the increase in incoming as well as outgoing traffic when avatars are used to enhance the presence as well as the (id)entity of a person. Here we explain the idea behind this picture. Our calculation method here is somewhat crude. However it serves well our purpose of rough estimation of the traffic increase due to virtual mobility and the presence of avatars.

† Sergey Shumsky[5] proposes an interesting self-organising semantic network, in which personal agents act on behalf of their users, thereby forming a distributed search engine, where each node, though possesing only limited amount of local information, can handle global queries.

Figure 8 Traffic increase formulae



Assume the incoming traffic arrival rate for the person depicted above is λ (in units of traffic arrival, for example calls per day or hits per hour) and each call or session duration is 1. Symbol p is the probability of the avatar owner being present to receive the call. The total incoming traffic load is then simply λp. When avatars are used, there are three steps in communication:

1 There will be a communication with an avatar first (in our case we assume avatars take all incoming traffic first before forwarding the call or informing the person). This communication is of duration α. This communication will probably be shorter in comparison to the length of the communication if the owner would have been present.
2 The avatar informs the person directly or in due time that there is (or there was) incoming traffic for him. We represent the duration of this communication with β.
3 The owner may decide to get in touch with the generator of incoming traffic. The duration of this communication session is represented by γ.

The total traffic generated till now is then $(\alpha + \beta + \gamma)\lambda$. Generally $\alpha + \beta + \gamma > 1$. Further we may assume that due to better reachability (popularity) of the person above, there will be an additional factor of traffic increase (factor c here).

Concerning the traffic triggered by the owner himself, assume that under normal conditions without avatars there is a traffic load of μ_0. When k avatars are used, each generates a traffic amount by itself: μ_1 to μ_k. Further, each avatar will probably communicate briefly to the owner the summary or results of its mission. For simplicity we have assumed here that this is a fraction ϵ of the traffic for all avatars. Finally, because avatars take over a part of activities from the owner, the owner himself would probably reduce his own traffic generation with factor $\delta < 1$. It is realistic to assume that the sum of the traffic generated in the case of avatars presence is (much) higher than μ_0.

A numeric example is given here with the following conservative assumptions: $p = 80\%$, $\alpha = 0.1$, $\beta = 0.1$, $\gamma = 0.9$, $c = 2.0$, $\delta = 0.2$, $\epsilon = 0.1$, $k = 3$, $\mu_1 = \mu_2 = \mu_3 = \mu_0$. With these values there will be 2.75 times increase in traffic triggered by others and 3.5 times traffic increase triggered by the owner. Please notice that this conservative estimation of traffic increase could in the future be multiplied by an order of magnitude proportional to the possible communication combinations shown in the (id)entity matrix of Figure 2.

Conclusions

Virtual mobility will:

- cause a landslide in society comparable with the shift from nomadic to colonist life – the pervasiveness of cyber life will dramatically change society and human behaviour this century;
- enable multidimensional life – people will travel and appear anywhere without carrying their atoms with them;
- be facilitated by human-like avatars;
- endanger the mental health of people who already struggle with their personality and identity. Future 'multimedia masks' will be far more sophisticated compared to today's text-based instant messaging hide and seek. Using mask avatars and the possibility to submerge oneself in cyber life will impoverish some people's real physical life.

Full-blown virtual mobility will cause an enormous new traffic flow:

- which is a new revenue generator for telcos and service providers;
- consisting of presence status information being sent to different groups of entities being people, avatars, applications running on several devices.

Market introduction of virtual mobility and avatars will be driven by:

- efficiency needs in general;
- the increasing value of time;
- cost and irritations concerning physical transport;
- higher oil price, saving energy and reducing global pollution (Kyoto).
- vanity or being uncertain about one's looks.

Avatars will:

- develop into an enormous diversity, some acting on behalf of people as their personal agents, others as agents representing organisations, animals, flora, precious objects – in cyberspace information will flow in a 6 × 6 identity matrix;
- communicate with people and with each other;

- be powered with artificial intelligence (Still simple in the short term, the ability to learn will arise. Human intuitive orientation skills are hard to learn and probably can not be matched so avatars will not become replicas of people, but will communicate in a human-like way with people and their appearance and behaviour will in some cases be anthropomorphic.);
- become more autonomous in the long term, take more complex decisions for their owner;
- organise themselves in hierarchical layers in the long term;
- not only be used for work, study or organisational purposes, but also for entertainment and psychological and emotional satisfaction;
- use sensors being their eyes and ears in cyberspace;
- require an *ergonomic* user interface/rules engine for the *avatar owner*. Note the 'rest of the world' is user of this avatar as well.

Avatars/active devices will outnumber people and dwarf the traffic volume between people.

References

1 Adams, W. H., Iyengar, G., Lin, C-Y., Naphade, M. R., Neti, C., Nock, H. J., and Smith, J. R. Semantic Indexing of Multimedia Content Using Visual, Audio and Text Cues. *Eurasip Journal on Applied Signal Processing*, Feb. 2003, **2003**(2).

2 Wayt Gibbs, W. Considerate Computing. *Scientific American*, Jan. 2005. pp. 40–46.

3 Gorniak, Peter and Roy, Deb. Speaking with your Sidekick: Understanding Situated Speech in Computer Role Playing Games. To appear in Proceedings of Artificial Intelligence and Digital Entertainment, 2005.

4 Moore, Gerhard. R., and Hobbs, D. J. An Experimental Study of the Effect of presence in Collaborative Virtual Environments. Proceedings of the British Computer Society Conference on Intelligent Agents, Mobile Media, and Internet Applications, 2001, Bradford, UK.

5 Shumsky, Sergey. Self-organizing Internet Semantic Network. White paper, by NeurOK, LLC, 2001.

6 Thrun, et al. Probabilistic Algorithms and the Interactive Museum Tour-Guide Robot Minerva. *The International Journal of Robotics Research*, 2000, **19**, pp. 972–999.

7 We thank Siemens for providing the basis for Figure 1 – Future image.

Biographies

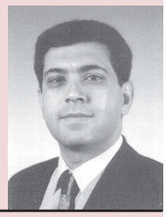
Nico Baken
Delft University of Technology and KPN

Professor Baken currently holds a part-time chair in the Telecommunications Department at Delft University of Technology alongside his primary position as Chief Architect for The Royal KPN, the Dutch incumbent operator in the Netherlands. His main interest concerns broadband networks and services, dealing with a broad range of aspects such as fibre access infrastructures, fixed-mobile convergence, services, operations, financial strategies such as the real option analysis, managerial complexity and regulations. Given this spectrum, he has been asked to advice the Dutch government on the matter of broadband and the roll out of fibre to the home in the national expert group broadband and in the Andriessen committee (former minister of Economic Affairs) to deal with the FTTH for Amsterdam and the Hague.

He finished Gymnasium β in 1973 and graduated, cum laude, in mathematics at Eindhoven University of Technology in 1981. He has published over 30 papers, holds several patents and won several prizes for his scientific work; for example, the Dr Neher Laboratory prize (yearly prize for the most outstanding researcher). He received his Ph.D. Thesis from the Delft University of Technology at the department of Electrical Engineering, working with Professor H. Blok and Professor A.T. de Hoop.

Edgar van Boven
KPN

Edgar van Boven studied electronics and IT at the Technical Highschool in Vlissingen. Though tempted to start an adventurous life as a jazz pianist, he graduated in 1987. After military service as a sergeant in a telecommunications battalion, he entered KPN. Until today, public telephony dominated his career from various viewpoints starting with hardware and software engineering, via operational network planning to architecture and programme management. In the late 1990's, he started to work on the evolution to voice over packet in the former Unisource Business Networks environment within KPN. Since 2001, he has also been guest lecturer at the Delft University of Technology. Currently, Edgar is working in the area of fixed mobile convergence and service architecture.



Ramin Hekmat
Delft University of Technology

Ramin Hekmat received his M.Sc. degree in electrical engineering from Delft University of Technology in The Netherlands in 1990. Since then he has worked for several telecommunication companies in The Netherlands and the United States in research and development as well as managerial positions. In September 2001, he started his Ph.D. work related to ad-hoc networking at TU Delft. His prime research interest includes multi-user communication systems, wireless communications and peer-to-peer networks.

Ludmila Menert
Consultant

Ludmila Menert graduated in General Linguistics and Phonetics at Leiden University, and in 1994 received her Ph.D. in Phonology and Phonetics from Utrecht University. After three years of teaching at the phonetics department of Utrecht University she made a career switch by co-founding Fluency Speech Technology. Within this company she co-authored Fluency, the best text-to-speech software for Dutch at that time. When the company was sold to Van Dale Lexicografie (Dutch dictionaries publisher), she joined Van Dale as a speech technology expert. Due to her involvement during the past years in a whole range of projects in the field of man–machine communication for her last employer, Comsys BV, her expertise now covers the broad area of deployment of language and speech technology in automated systems as well as usability of speech applications. Currently she works as an independent consultant.