Creating virtual objects

Antti Pirhonen
pianta@cc.jyu.fi
University of Jyväskylä

Abstract

User-interface is much more than just a sum of its parts. Therefore, the primary task of a user-interface designer is not to design a set of presentation elements but to design meaningful wholes. The focus of the current paper is in the analysis of the relationships among presentation elements. The underlying objective is to support the design of highly integrated multimodal combinations of presentation elements. The discussion is based on human interaction with real-life objects.

Keywords: interface design, information presentation, multimedia
BRT Keywords: FB0403, AD03, HD01, CA14

Introduction

Recent efforts in the research of multimodal user-interfaces have a pragmatic background. In the past, an average user of a computer was an expert in technology. The focus was in the primitive functions of the processor. The results of the calculus performed by the computer were often presented in a form that could be understood only by an expert. Nowadays, microcomputers have spread all over the community. The newest information technology is used by all kinds of people for numerous kinds of tasks. An average user of a computer is not interested in the technological details and functions of a microprocessor. Instead, computer is used as a tool that is supposed to perform a task – the user is not interested in *how* it is performed. Therefore, the quality of the user-interface determines, to high extent, the quality of the system.

The quality of the user-interface is an extremely ambiguous concept. The criteria for it depend, above all, on the paradigm within which the quality is assessed. Even as widely used criterion as *usability* contains so much vague features that it does not work as a guideline in practical UI-design. The problem in the articulation of criteria for user-interface is that when trying to create concepts that are independent of single disciplines (or interdisciplinary), the paradigms of single disciplines cannot be used in the definition. Therefore, when striving toward explicit concepts the basis has to be an existing discipline with its paradigms.

The basic principles of the proposed approach are presented in the third chapter, in which the experiencing of a virtual object is compared with the experiencing of a real-life object. The use of this metaphor in design requires proper understanding of concepts like modality and the concepts that describe the relationships among presentation elements. Therefore, succeeding chapters contain analysis of those concepts. The last two chapters discuss the roles of the designer and the user in their communication with the help of virtual objects.

In search for an alternative approach to UI-design: focusing on the relationships among single elements

In the current work, the central paradigms when discussing UI-design are related to communication – design is understood as creation of messages for the user. In the approach, an atomic unit is a *message*. In physical level, the counterpart of a message is a *presentation element,* not some technical unit like pixel. Therefore, the object of design can be seen either a message or a presentation element, according to whether the emphasis is in the meaning or the appearance of the element. But the fundamental task of a designer is not to create a set of presentation elements but to create a user-interface, which is certainly much more: First, presentation elements are usually only output elements, and input elements are at least as central in user-interface. Second, user-interface should be a meaningful whole, not an arbitrary collection of independent details. In other words, the output of user-interface is more than a sum of presentation elements. Therefore, the focus of a designer should be, not only in the design of presentation elements, but also in the relationships among them. This would help the designer in two kinds of tasks:

1. To control the whole user-interface. In the analysis of the user-interface the whole has to be first split into details. If the only level of details were the one of presentation elements, presentation elements would be handled isolated from the rest of the user-interface. Therefore, if the level of relationships were taken into account besides the level of presentation elements, the elements would be seen, not only by themselves, but also as a part of the whole.
2. To create *virtual objects* instead of single presentation elements. This means here an attempt to create multimodal combinations of presentation elements in which different elements are strongly integrated.

Virtual object vs. real-life object

A multimodal communication environment makes it possible to create multimodal combinations of messages that can be interpreted to form a single whole. But should this whole be called a message or a combination of messages? This problem concretizes the core of the nature of creating virtual environments (Fisher, 1990) consisting of virtual objects. The comparison between a real-world object and a virtual object is illustrated in figure 1. The drawings illustrate the formation of a mental representation on the basis of external and internal events. In the first case, the physical origin of the process that results in the formation of mental representation is a single physical object. That object affects its environment in multiple ways. For example:

- It absorbs part of the light that falls on it and reflects the rest.
- If it or some part of it moves, it causes changes in the material that surrounds it.
- It fills an identifiable portion of space.

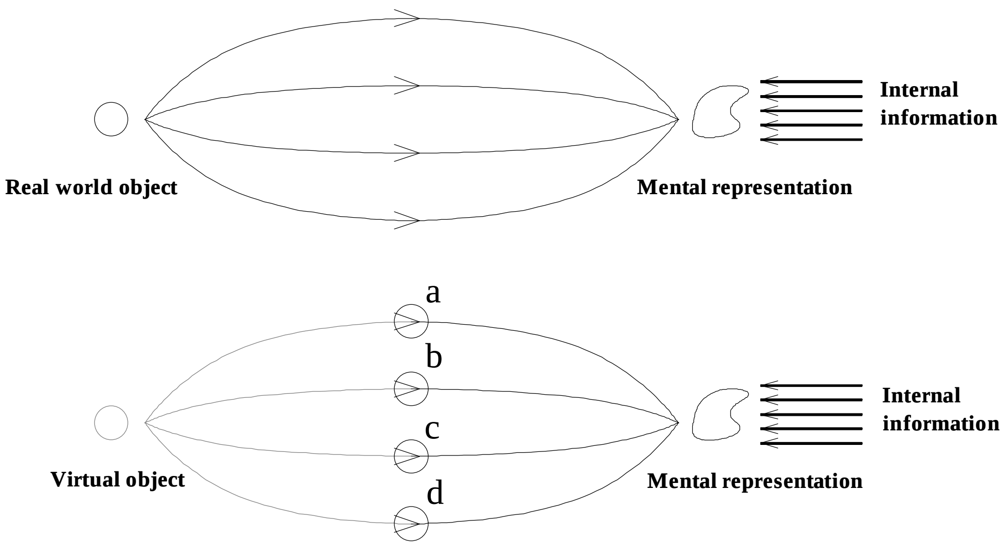

Figure 1: Formation of mental representation on the basis of real and virtual object.

- It might emit particles in its environment.

From the point of view of a human being, the first property causes a visual perception. The second, assuming that the movement is vibration and the frequency is between 20 and 20000 Hz, produces an audio perception. Because of the third property it is possible to identify it by touching it. The fourth property mentioned might cause, in certain conditions, a perception of smell.

The physical effects of the object in its environment are illustrated in the drawing with four arrows. These external information sources, together with existing mental structures (in the drawing labeled as internal information sources), cause a mental representation of the object.

The lower drawing illustrates a situation in which a similar mental representation results but the physical object is missing. In this case, the primary physical origin is not one single object but several qualitatively different presentation elements (a, b, c, and d). Each of these presentation elements must be produced individually. The physical property that combines them is that they are, for example, activated simultaneously as a result of the same user action.

This difference between a real and a virtual object is important to be understood by a designer. Since a virtual object is not naturally multimodal like a real life object is, the nature of the virtual object has to be individually considered from the point-of-view of each available modality.

Concept of modality in multimodal information presentation

The word modality usually refers to sensory systems. In the current microcomputer environment, multimodal information presentation means that both visual (via monitor) and audio information (via speakers) is provided. The division between audio and visual information is clear and easy to make. Especially in a multimedia environment, this division seems natural since the same classification refers directly to the two most common presentation devices, visual display units and sound devices (loudspeakers or headphones connected to a sound card). But when discussing human information processing and multimodal virtual objects, the distinction between perceptual systems is far too crude to provide a basis for analysis of the consequences of message format selection. Therefore, a much more fine-grained classification is necessary.

A typical further division made after the division into audio and visual information is to divide both audio and visual classes into verbal and non-verbal sub-classes (e.g., Mayes, 1992). Allan Paivio (1986, p. 57) uses this kind of classification, enhanced with haptic, taste, and smell sensory systems (table 1). Since taste and smell have neither verbal properties nor presentation devices in a contemporary multimedia workstation, they are skipped here. Neither is the haptic system taken into account, as the focus of the current work is in information presentation. Braden (1992) used the same simple classification when illustrating different possible ways to combine two presentation elements in an audio-visual presentation (figure 2).

Table 1: The classification of Paivio (1986, p. 57) with examples. (Brackets and italics added by author.)

	Symbolic Systems	
Sensorimotor	Verbal	Nonverbal
Visual	Visual words	Visual objects
Auditory	Auditory words	Environmental sounds
(Haptic	*Writing patterns*	*"Feel" of objects*
Taste	-	*Taste memories*
Smell	-	*Olfactory memories)*

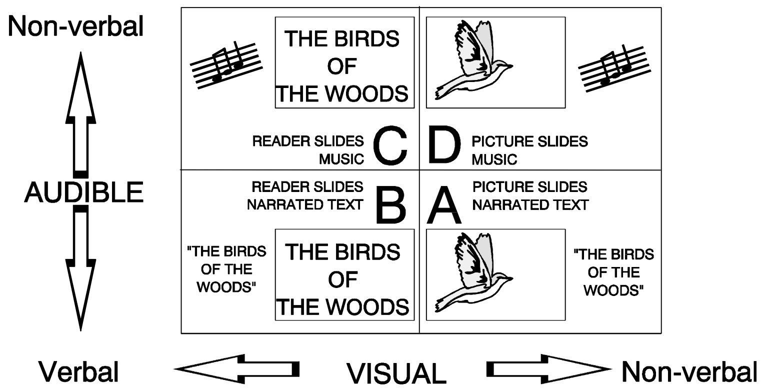

Figure 2: A 2x2 matrix concerning the classification of combined audio-visual messages. (Braden, 1992)

Bernsen (1995) takes a long step forward in the classification of output modalities. In modality theory, Bernsen suggests a division of the concept of modality into sensory and representational modalities. By sensory modalities he means the division into modalities according to the sensory system applied, such as hearing and vision. By representational modality he refers to an elaborate classification of qualitatively different kinds of information forms. In his theory, different forms of presentation elements are classified according to whether they are linguistic, analogous with their referent, arbitrary, static or dynamic, and whether they are physically graphics, sound, or touchable information. The classification is coherent with the more common classification that only takes into account the sensory system and whether an information unit is verbal or not. However, modality theory goes much further, resulting in 28 different classes. For example, in Bernsen’s classification, a case in which a text unit is in one location on a VDU (visual display unit) differs from a case in which the same text unit is in motion from one location to another because they differ with respect of their static/dynamic nature. Modality theory with its classification contributes the shift of focus from physical media to human information processing because each medium is used to present information in several representational modalities. For example, all ten graphical modalities are presented with a VDU.

When discussing information presentation to humans, it is obviously essential that the focus be on human mental processes instead of technical implementations. However, an astonishingly high number of research papers has been published concerning the effects of multimedia in, for example, learning.[1] This observation is made and strongly criticized by Clark and Craig (1992), who recommend a shift of attention in learning studies from technology used to instructional methods. For the same reason, in the present work different forms of the word “multimodal” are used instead of “multimedia” even when referring to applications that are run on multimedia workstations.[2] Thus, the conceptual basis of the present work is independent of the technology used. Additionally, since it is possible to present information via one sensory modality in several qualitatively different ways that require quite different kinds of cognitive processing, “modality” and related concepts refer here to representational modalities. Thus, multimodal presentation, according to the current definition, may be designed to apply to one single presentation device. For example, a combination of a narration and background music fulfills the criteria for multimodality.

Rationale: why to use multiple modalities

Using multiple modalities when creating virtual objects is complicated and causes several problems that have to be solved. Therefore, there must be a good reason for going in for these efforts. Two possible reasons are naturalness and an attempt to provide much information:

---

[1] A typical case is the approach of Najjar (1996). The conclusions about the “effects of multimedia” in learning are based on a comparison between a lecture and computer-based instruction.

[2] The conception of multimedia and the related concepts is based on the broad reviews of Galbreath (1992) and Tolhurst (1995) that clearly suggest limiting the use of the word “multimedia” to a technological context. Dictionaries of the domain suggest a similar definition (e.g., Latham, 1995).

1. Naturalness: An ambitious designer of a virtual object would probably wish to use all available media in order to provide as natural an impression as possible. Since we usually experience the properties of real-life objects in a multimodal nature, the challenge of the designer can be seen to support this kind of interaction (the ecological approach, Gibson, 1979). The primary motivation to naturalness might be, e.g., user satisfaction. Also, this kind of approach utilizes the existing interaction skills. Instead of forcing the user to learn some new strategy in human computer interaction, the strategies we are already able to use are exploited in a computerized environment. This exploitation of "natural" strategies also includes the existing skills in interacting with, e.g., technology. For example, a push-button metaphor of graphical user-interfaces utilizes the convention of a switch of an electronic circuit.
2. By using multiple modalities, more information can be provided than by one modality. By using sounds, it is possible to present such information about a dog that would be impossible to present via other modalities. Actually, each modality has its unique features that make it, at least to some extent, irreplaceable. On the other hand, multiple modalities can be used in order to utilize human information processing capacity. In some cases, it is even possible to process information of several independent presentation elements without any interference.

Virtual objects as combinations of presentation elements of different modalities

The two rationales above may result in quite different design decisions. If the uppermost aim is to provide as much information as possible (the second rationale), the creation of a meaningful whole will very likely be seen as a puzzle. For example, one presentation element conveys information about the appearance of a dog; the other is about its voice. Different presentation elements thus complement each other. Overlapping of contents of different presentation elements is either unnecessary or even avoided. The view of information is the one of information theory (Shannon & Weaver, 1949): information that is presented more than once is not information at all. The ideal combination of presentation elements could be illustrated as in figure 3.

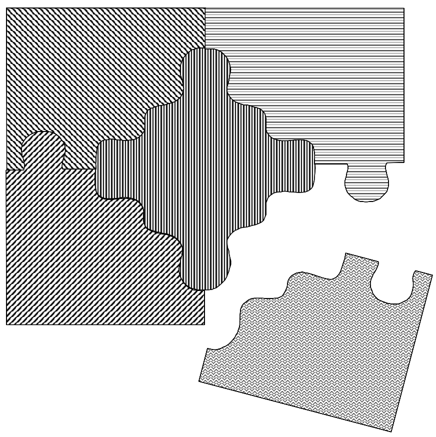

Figure 3: A puzzle model: presentation elements complement each other.

However, creating a "puzzle" is problematic. This kind of neat combination means that messages conveyed by presentation elements can be analyzed in detail. Detailed analysis of messages is, in turn, impossible for two reasons:

1. The essence of a meaning of a message is frequently beyond the reach of human language. Especially earlier experiences about a subject (in figure 1: internal information) are to high extent non-verbal in nature. For example, values and affections related to the subject are impossible or hard to verbalize.
2. The formation of meaning is to high extent subjective process. Therefore, detailed assumptions about the meanings evoked by a presentation element reject the fact of unique, individual interpretations.

Instead of a neat puzzle, the result is merely something like the drawing in figure 4. Different elements convey meanings that are partly unique, partly redundant. There might even be content that is present in all presentation elements (intersections in figure 4).

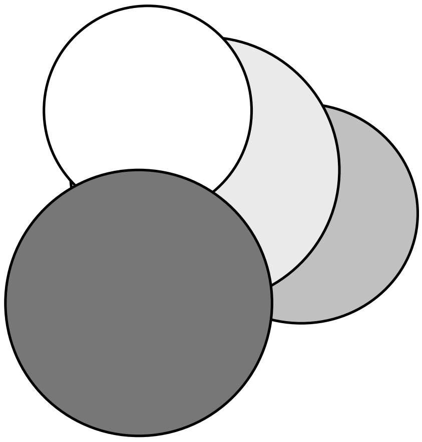

Figure 4: A combination of presentation elements with overlapping contents.

For the needs of describing the relationships among presentation elements, a conceptual tool has recently been developed (Pirhonen, 1998b). In this model, the relationships are classified according to the amount of information in common in presentation elements. The basic dimension of the model is defined by two extremes, *redundant* and *distinct* cases. In figure 5, the first and the last row represent the extreme ends of the dimension, and the second and the third rows different intermediate cases.

The essence of the drawing is to illustrate the change in the quantity of information in common between two messages. While in *distinct* conditions the messages have nothing to do with each other, in *redundant* conditions they concern exactly the same entities. In practice, most cases can be located somewhere between these extremes. That is why the intermediate area has to be elaborated.

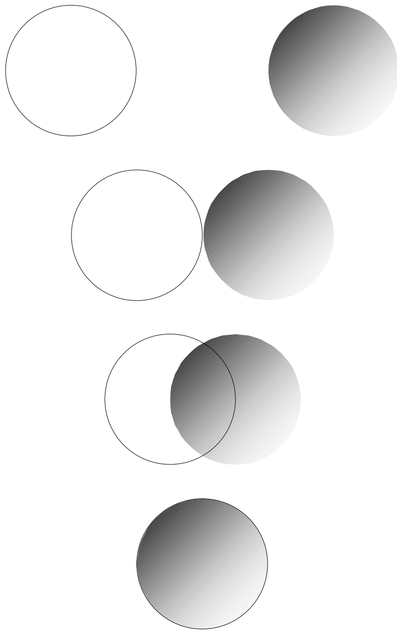

Figure 5: Possible relationships between two elements. (Pirhonen, 1998a)

In the second row, the circles are tangents to each other and thus, have only one point in common. But they still form a figure together that is no longer two circles but a single whole. This figure represents a case in which the messages are about the same object, but contain quite different information about it.

In addition, the case in which two elements complement each other, is elaborated in the model, resulting in a two dimensional model (figure 6). In the model, a relationship is analyzed according to the amount of information in common (horizontal axis) and according to how the relationship fits our conventions

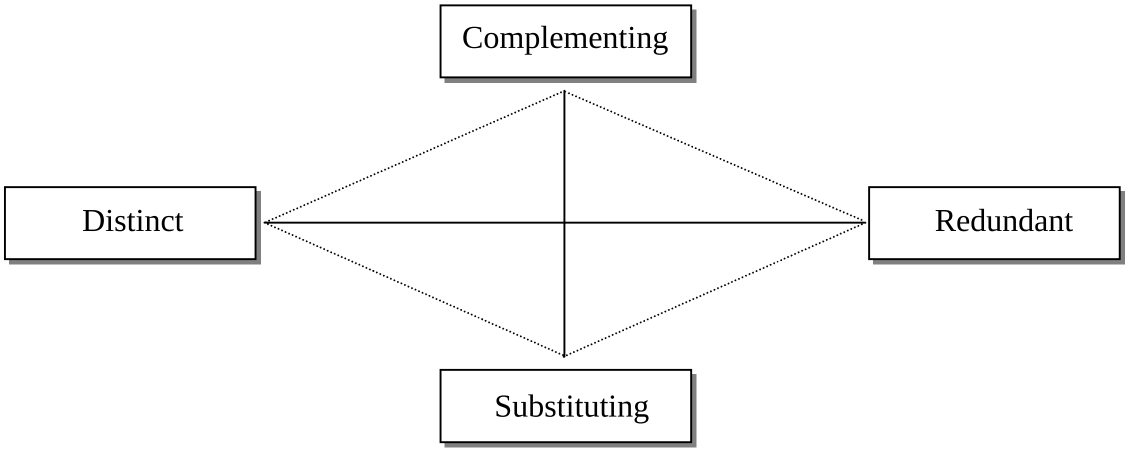

Figure 6: Different relationship types. (Pirhonen, 1998a)

(vertical axis).

The model works as a conceptual tool when designing multimodal combinations of presentation elements. With it, most relationships can be described. The model thus helps to control the whole by providing the designer a way to figure out the relationships and to communicate with the project group.

Qualities behind the facts

The model above was originally meant to cover all essential qualities of the relationships among different presentation elements of a user-interface. However, when it was tested in user-interface analysis (Pirhonen, 1998b), it was found that perhaps the most important qualities of the relationships are difficult or impossible to be verbalized. In this kind of case, the relationship could be located in the dimensions of the model, but it proved problematic to verbalize the reasons for the localization decision. For example, a combination of a high-quality picture of a trumpet and a simultaneous sound sample was argued to be redundant, but no reason could be mentioned. It was simply *experienced*, using the words of the subjects, “something more than just complementing”. This kind of observations force to consider, whether the model should be further developed to cover more levels than just facts (that are easy to verbalize).

An attempt to go beyond logical and physical facts is extremely problematic and leads the discussion inevitably to the fundamental epistemological and ontological questions. Science is traditionally based on verbal argumentation. Therefore, an attempt to reach dimensions beyond perceptible facts may question the credibility of the approach. However, in certain cases the verbalized interpretations and other reactions evoked by a virtual object may differ so much that expanding the model cannot be avoided. Using the same example again, the difference of interpretations is clear when discussing a virtual dog. In the creation of a virtual dog, the designer might wish to deliver information about the appearance and the voice of a dog. In factual level, the picture of a dog contains information about the physical properties of one dog. Respectively, the sound sample provides information about the sound properties of dogs barking. Mechanically seen, the message of the picture would have a verbal interpretation like “a dog stands with four feet, is brown in color, has a tail,…”. In other words, only currently observable information is taken into account. The sound sample could be interpreted in the same manner, by describing the audible properties. The picture and the sound sample thereby complement each other; the relationship between these two presentation elements would be classified as strongly *complementing*. However, since most people have plenty of experience about dogs, the role of earlier experience is significant when discussing the message. For example, the actual message of a picture of a dog would be “Run away!” for a person who has been frightened by a dog and is therefore afraid of them. The sound sample would probably have just the same effect on the person. Therefore, even if in the factual level the relationship was complementing, in another level it is clearly *redundant* since both presentation elements evoked similar affections.

Now it seems that the content of messages vary markedly according to the level in which the messages are interpreted. Thus far, the “level” is referred to in a loose manner, on the basis of a couple of examples like “the level of facts” or speaking about “affections.” A more exact view of different levels of interpretation is clearly needed in order to elaborate the redundancy model.

The need to analyze real-life objects in a deep manner, i.e., the need to go beyond the physical and logical facts, has resulted in different models in different disciplines. In education, learning objectives have traditionally been split into cognitive, affective, and psychomotor levels. Bloom (1972) developed this taxonomy for the needs of planning and evaluating education. It is a clear example of an attempt to broaden the conceptual basis of the analysis of the interaction between a human being and his environment. Bloom's taxonomy can be seen as a rough outlining of the third dimension of the proposed model (figure 5), or at least an attempt to name more levels than just facts. The attempt is respectable especially since it is done in the behaviouristic paradigm. Thus, only perspectives that can be seen to have solid connections to overt behaviour could be taken into account.

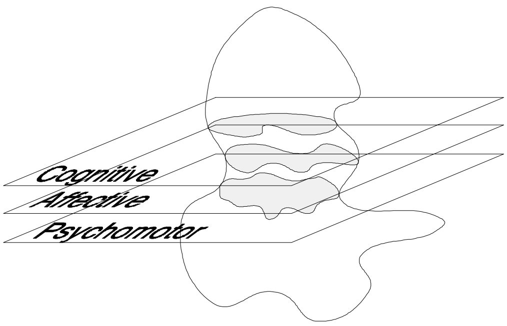

Figure 7: Three layers derived from Bloom's Taxonomy.

Figure 7 illustrates the current view of the contribution of Bloom's taxonomy to analysis of a real-life entity at more than one level. By identifying and naming three levels the model is a much better basis for designing a presentation of the entity than a description concerning just facts (the nearest counterpart of facts in the figure is the cognitive level.) The figure illustrates how the taxonomy, however, only concerns some thin slices or sections of the entity. By increasing the number of layers it is possible to get a better and better view of the whole. But finally, that approach can not lead very far since basically it is a question of classification, and classification inevitably means reduction. If it were possible to define a whole new perspective, we could see the whole profile of the object instead of a set of sections. The new perspective would mean the definition of the third dimension discussed above.

As mentioned, Bloom's model is constrained by its behaviouristic basis.

Conscious experience, understood in the way it is handled in consciousness studies (e.g., Chalmers, 1996) could be a promising framework in formulating the new perspective and dimension within the contemporary paradigms. Likewise, the implicational level in the ICS-model (Interacting Cognitive Subsystems, Barnard & May, 1993), if further developed, could be the new perspective. The problem of using the current version of the ICS-model is that the properties of the implicational level are not yet fully articulated. Or the discussion about the dimension could lead to analysis about the relationship between science and art and their roles in outlining reality. However, the need to expand the two-dimensional model is so apparent that this has to be considered even if the expansion is likely to make the model much more complicated.

Although it is not yet possible to define a third dimension in the model of relationships between presentation elements, the contribution of the approach to the discussion about redundancy can be illustrated as in figure 8. The figure illustrates two objects and the multiple levels that define their properties. The objects illustrate mental representations of concrete or abstract entities. The tiny white area on the top of each refers to the layer of factual information. The illustration stresses the conception according to which this layer is only the tip of an iceberg. In the illustrated case, the relationship would be classified as *distinct* if only the uppermost layer is taken into consideration. But if the interpretation of the underlying mental representations goes deeper, clear similarities are found at some level or levels. Some degree of redundancy takes place in a form that cannot be explicitly defined or described verbally. If the real-life objects that have evoked the representations illustrated in figure 8 are two presentation elements, they can be seen, according to the illustration, to form a meaningful whole. That whole is – in turn – a representation of a virtual object.

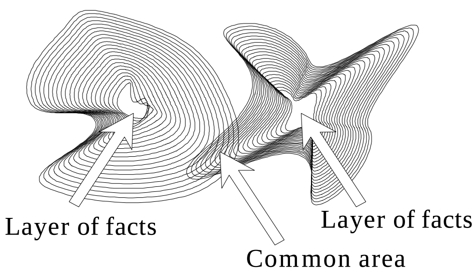

Figure 8: Two representations with common area outside the layer of facts. (Pirhonen, 1997)

The example above about a virtual object was a dog. Creating a virtual dog is, in a certain sense, quite easy; the designer tries to imitate a real-life object. The process of creating such a virtual object can be seen as a mechanical reproduction of reality. The designer takes a photo of a dog and records its bark. The material is then simply stored in digital form and linked or embedded to an appropriate application. In fact, the process is, of course, much more complicated. Just the taking of a simple photograph of a dog implies an enormous number of decisions that all affect the result. A designer's skills in using a picture as a communication tool define the quality of the result.

In the illustration (figure 1), cases that were compared with each other were the formation of a mental representation of a physical real-life object and the formation of a mental representation on the basis of a virtual object. But if the virtual object lacks a real-life counterpart, there is no physical object that could be imitated or that could be reproduced. In that kind of case, the origin of the presentation is the designer's mental representation of the object. It is a question of the skills of the designer to express with the available means something that has a highly abstract and subjective form in her or his mind. On the other hand, this can also be generalized to cover the cases in which a virtual object does have a concrete counterpart. Finally, the designer's mental representation of the object rules her or his decisions in the design process. Referring to the example above about the complexity of taking a photo, the mental representation of the object of a photo rules the decisions in that situation. The whole repertoire of photographic means of expression is – or at least should be – in the use of the designer to express something essential about her or his mental object whose physical counterpart is called "a dog."

There is a danger that the illustration (figure 1) may be interpreted as a slightly modified version of the mathematical theory of communication. It has to be admitted that it illustrates transference of information with its one-way arrows. The arrows cannot even be two-headed since there is usually no interaction between the designer and the user. But the essential difference between this illustration and the process models of communication is that it is not a question of an attempt to copy a meaning from the consciousness of the sender (designer) to the consciousness of the recipient (user). The illustration and the underlying idea emphasize the role of the activity of the user when working with external information on the basis of internal information. While in process models human qualities are handled as constraints, in this model they are a vital resource. From the perspective of process models it would be a great success in communication if we someday had direct access to the physical structures that store mental representations. From the point of view presented here, it would mean skipping the substance of human communication, active interpretation.

Design process as communication

Finally, the essence of the process of presenting information in a multimodal format seems to be quite analogous with everyday communication between human beings. Basically, however sophisticated the information technology used to deliver information is, a multimedia product is a collection of messages from a designer to a user. In other words, it is a question of mediated human-human communication. Both *how* and *what* information is presented inevitably reflects the mental life of a designer. Technical devices are tools that enable a designer to express her or himself. The skills in utilizing the available technology are certainly of great importance. But in a more central role are general communication skills. How could a designer who does clumsy writing with a typewriter or with a pen, or always becomes misunderstood because of his shortcomings in verbal communication, write fluent and understandable text with a multimedia development application? Or how could a designer who took totally unimaginative and boring photos on his family's holiday in Hawaii suddenly take fascinating and inspiriting photos only because they are taken for a multimedia product?

New information technology contributes to communication by providing a powerful way to present huge amounts of information. The central question is, for what reason is the information presented. There are at least three possibilities:

1. Information is presented for the senses of a user. A sound device is used because the user has ears. A VDU is used because of the existence of eyes. The amount and the format of presented information depend on the content of the presentation and

the capacity of the devices. Technical expertise is central.

2. Information is presented for the cognition of a user. Detailed information about the cognitive system is utilized in order to make the presentation effective. The amount of information is rationed out and encoded on the basis of the cognitive models used.
3. Information is presented for the consciousness of a user. Information presentation is seen as a challenge to express meanings, many of which cannot be directly presented with the devices available. The active, interpretative role of the user is therefore essential. Skills in expressing oneself and communication are required.

The classification above should not be seen as a trivial hierarchy of the values of different approaches. Each of these may result in failure or success. It may even be difficult to say on the basis of a completed application which – if any – of these approaches has been applied. For example, it is possible that the production has been totally technology driven (approach 1), but the content has fit so well in the technology available that the result is both cognitively ergonomic (approach 2) and evokes mental activity (approach 3). Of course, in this kind of case the success has resulted accidentally – despite the technical bias – and would therefore not work as a pattern to design. On the other hand, the approach to design may be extremely ambitious, in which many subtle nuances of human qualities are taken into account, but still end up to a failure: however much e.g. artistic resources are paid in design in order to reach the consciousness of the user (approach 3), they would not compensate possible lack of expertise in cognition (approach 2) or technology (approach 3).

These three approaches should be seen to complement each other. Each of them should be taken into consideration in order to design a high-quality multimodal user-interface. The more technical (approach 1) and psychological (approach 2) expertise as well as creativity and communication skills (approach 3) available in a multimedia project, the better in order to communicate with the user with the help of virtual objects.

The user as an active knowledge constructor

In figure 1, the shapes that signify real world or virtual objects and the resulting mental representations differ from each other. In the early process models of communication this would have been interpreted as a result of noise – disorder or interference in the communication channel. In those models, the target was to reproduce the original meaning. The activity of the receiver meant, in that view, preparedness to receive information units. In the current view, it is admitted that meaning can never be transferred from the consciousness of one person to the consciousness of another person. This is not a constraint or an undesirable fact with which we simply have to cope. Rather, this illustrates the creativity and activity of a human being. The contribution of the current work to design is therefore *not* to support *effective information transfer*. Instead, the target is to *present information in a form that is applicable for the users' needs*. There is an important difference between these two approaches. In the first one, the purpose of the communication via multimedia is to influence and control; multimedia presentation is expected to cause more or less permanent changes in the mental structures of the user.

The expertise on human mental life is used to perform these changes effectively. In this approach, redundancy, in the sense it is presented in the current work, is one way of having a stronger effect on the user. On the contrary, the second approach leads to a more user-centered view. Information is *provided* for the user, who either uses it or not in his or her constant knowledge construction process.

The approach that stresses the activity of the user might result in overflowing user-interfaces. As much information as possible is presented, and the user is expected to choose the appropriate pieces. The approach in which the user is a passive receiver of information, may, in turn, result in much clearer user-interfaces since the attention of the user is then easier to control. However, it should not be thought that this is necessarily the case and that the relying on the activity of the user results in confused user-interfaces. If the user-interface is filled up with elements, can we assume that all the information that is coded in those elements is really presented for the user? Referring to the classification above, it can be said that overflowing user-interface information is presented to the senses of the user. As this is seldom enough, we have to turn attention from the quantity of presentation elements to their quality. The designer should not assume that a certain piece of information is presented for the user when it is coded as a user-interface element since other elements can drown it out. The key is to piece the whole user-interface together and to analyze each single element in relation to each other and the whole.

References

Bernsen, N. O. 1995. A toolbox of output modalities: Representing output information in multimodal interfaces. Esprit Basic Research Action 7040: The Amodeus Project, document TM/WP21.

Bloom, B. S. 1972. Taxonomy of educational objectives: The classification of educational goals. New York: McKay.

Braden, R. A. 1992. Examining Visual Verbal Relationships. In Art, science & visual literacy: Selected readings from the annual conference of the international visual literacy association (24th, Pittsburgh, PA, September 30 - October 4, 1992), 229 - 238. (Eric Document Reproduction Service No. ED 373 717)

Chalmers, D. J. 1996. The conscious mind. In search of a fundamental theory. New York: Oxford University Press.

Clark, R. E. & Craig, T. G. 1992. Research and theory on multi-media learning effects. In M. Giardina (Ed.), Interactive multimedia learning environments: human factors and technical considerations on design issues. NATO ASI Series F: Computer and System Sciences, 93. Heidelberg: Springer-Verlag, 19 - 30.

Fisher, S. S. 1990. Virtual interface environments. In B. Laurel (Ed.), The art of human computer interface design. Reading, MA: Addison Wesley, 423 - 438.

Galbreath, J. 1992. The educational buzzword of the 1990's: Multimedia, or is it hypermedia, or interactive multimedia, or...? Educational Technology 32(4), 15 - 19.

Gibson, J. J. 1979. The ecological approach to visual perception. Boston: Houghton Mifflin.

Latham, R. 1995. The dictionary of computer graphics and virtual reality. New York: Springer-Verlag.

Mayes, J. T. 1992. The 'M-Word'. In A. D. N. Edwards & S. Holland (Eds.), Multimedia interface design in education. NATO ASI Series F: Computer and system sciences, 76. Heidelberg: Springer-Verlag, 1 - 22.

Najjar, L. J. 1996. Multimedia information and learning. Journal of Educational Multimedia and Hypermedia 5, 129 - 150.

Paivio, A. 1986. Mental representations. A dual coding approach. New York: Oxford University Press.

Pirhonen, A. 1997. Meaningful whole or a jumble? The problem of designing a combination of output elements. Management 2(8), 62 - 65.

Pirhonen, A. 1998(a). Analysis of the concept of redundancy concerning the design of multimodal combinations of output-elements. In: Werner, B. APCHI 98. Proceedings in the 3rd Asia Pacific Computer Human Interaction, July 15 - 17 1998, Shonan Village Center, Japan. Los Alamitos (CA): IEEE, 273 - 278.

Pirhonen, A. 1998(b). Redundancy as a criterion for multimodal user-interfaces. Jyväskylä, University of Jyväskylä. Jyväskylä studies in education, psychology and social research 147. (Diss.)

Shannon, C. E. & Weaver, W. 1949 / 1964. The mathematical theory of communication. Urbana: The University of Illinois Press.

Tolhurst, D. 1995. Hypertext, hypermedia, multimedia defined. Educational Technology 35(2), 21 - 26.