The AMODEUS Project

ESPRIT Basic ResearchAction 7040

M1.9 Report
S4: Visual Presentations
S5: Classification of Modalities

Daniel Salber
Laboratoire de Génie Informatique, University of Grenoble

25 May 1994

Amodeus Project Document: IR/PRES8

AMODEUS Partners:
MRC Applied Psychology Unit, Cambridge, UK (APU)
Depts of Computer Science & Psychology, University of York, UK. (YORK)
Laboratoire de Genie Informatique, University of Grenoble, France.(LGI)
Department of Psychology, University of Copenhagen, Denmark. (CUP)
Dept. of Computer & Information ScienceLinköping University, S. (IDA)
Dept. of Mathematics, University of the Aegean Greece (UoA)
Centre for Cognitive Informatics, Roskilde, Denmark(CCI)
Rank Xerox EuroPARC, Cambridge, UK.(RXEP)
CNR CNUCE, Pisa Italy (CNR,CNUCE)

S4: Visual Presentations

Sophie SCOTT, MRC-APU: Structural Descriptions

We actively perceive information about the world. Perception is a two-way process: getting information and bringing some.
Visual structure can be described as a hierarchy when zooming in on the structure (e.g., office, far wall, desk, pad, text, word, line). The structure defines the way we zoom in (constrains the zooming). But it doesn't imply that we start at the top level to make sense of the whole. When we're looking for an object in a room, the structure guides the search (navigation). Both top-down and bottom-up processing.
The psychological subject is part of the structure.
Rules of visual structure navigation: possible movements;

- go through the intermediate levels,
- move within a level,
- movements between and within levels are orthogonal (no jumps)

Structure of items of the world: what makes an object
features (shared): location, color, hue, proximity, texture, boundaries, junctions, uniform connectedness (e.g., glass), emergent properties
What makes a psychological subject? *the focus of attention*
Properties (e.g., boundaries) can be ambiguous (e.g., the pelican/rabbit picture)
Features such as hue help and influence the perception of the structure.
Size and length help also in a different way (capture attention).
Some attributes may override each other: e.g., color >> form
Conclusion:
perception is an active process of structuring the world
hierarchy
psychological subject depends on the attributes of the visual structure
Q&As:
• Structure has different levels of granularity. "Jumps" are done along the hierarchy. (guided by predicates+principles)
• similar to guidelines: they capture a lot of truth but are not sufficient
• problem with transfer (RP4)
• Ambiguity of the term predicate:
Example: pen as psychological subject: "next to bottle" could be the predicate.
Cross on line icon: cross is the psychological subject and the predicate is "superimposed on a line". But with a document icon, the icon boundary is the psychological subject

David DUKE, YORK CS: A Theory of Presentations

Deals with the structure of presentation, within a system model. Aim is to prescribe the structure of the presentation.
General framework:

- percepts: what can be perceived, e.g., a text string, a snapshot
- modality: how something is perceived, encompasses type of the percept, e.g., visual, audio, tactile,
- chunking: how percepts are structured, how the system intends the user to perceive.

Presentation is the combination of percepts, modality mapping and a chunking relationship, that is: a set—from a system point of view—of percepts is a percept, e.g., a scrollbar is a chunking of arrows and slider.
Every percept must be assigned a modality. For chunking, the set of percepts are part of the presentation as is the chunked percept. Modalities are combined.
Visual structure: percept occupies an extent of space on the display. Visual space has a structure (is topological). The extent of a percept part of a chunk is part of the extent of the chunk.
Refinement of the visual presentation: percepts can be organised into layers. Layered presentation is a combination of layered displays and overlapping percepts; a window is a self-contained presentation (frame + contents).
Time and audio: time is usually handled implicitly by formal methods (state #2 comes after the state #1, but this relation is implicit). Time is spread throughout the presentation.
Time itself is a modality. A timed percept has an extent in time. Each percept occupies both space and time: but unfiying space and time in the frame of formal methods is messy (i.e., complicated and difficult to work with).
Audio is a special case: alarms are atomic (that is they are either present or not present); a message (e.g., "mind the gap") is decomposed in smaller percepts and the system relies on *human* resources for chunking.
Question: The same human chunking happens for visual display?
Answer: it is part of the specification for visual presentation but audio is not specifiable in the same way.
Audio: to realize chunking, some percepts are buffered by the human (this is an assumption).
Summary:
This presentation model has been used in 3 analyses (Matis, AV Connections, CERD), although at a gross level (i.e., identify modality). It is not simple to do using Z or VDM.
Q&As:
Question: (suggestion) use of temporal logic.
Answer: gets very messy (attempted with Matis). Clashes with the standard approaches of formal methods.
Question: embodiment of psychological principles?
Answer: not the aim of system modelling. Attempt with work on formalisation of ICS, though. Have to draw the line between system modelling (what the system does) and user modelling (what the user perceives) (e.g., chunking by color: depends on the intent). More generally it is a problem of conformance, system modelling can say where the conformance will exist, user modelling will say where it is useful.

David DUCE, RAL: Visualisation

Initial motivation for USA visualization initiative: to present supercomputers data.
Dictionary definition: picture formed by visualizing.
Use of color to express structuring data. Geometric shapes.
Alternative term: "Vacualization": remove all meaning. :-)
Problem of distortion of presentation (honesty).
[Fred Brooks] collection of ad hoc techniques, not a real discipline.
Visualization ought to be designed. Honesty/truthfulness.

Frameworks:

- Haber&McNabb; three steps: data enrichment, mapping, rendering
- Bergeron: visualization can be descriptive, analytic or exploratory
- GRASPARC Model (see paper). math. modelling; three layers.

Telling the truth: from a table of data (oxygen = f(time)) -> plot; interest is between the points "continuity". But oxygen gets negative! -> use a method of interpolation (should preserve the properties of the data). So be careful with data enrichment.
Be careful too at the rendering level: e.g., display data sets using false color (default colormap is blue to red—not linear). Treinish at IBM Yorktown Heights is working on an architecture for rule-based visualisation which has some knowledge of perception and cognition and will suggest an appropriate visualisation technique to use based on properties of the data to be visualised.
Faithful representation vs. feature highlighting -> use color in a different way.
BOZ (see paper): way of describing a task + perceptual operator substitution + structuring
Example from UoPittsburgh: task: find connecting flights from Pittsburgh to Mexico city (see paper).
perceptual operators substitution/description/perceptual data structure -> the system suggests ways of presenting the info (table, different graphs highlighting different features of the data).
As a conclusion:
Knowledge of human perception is necessary for visualization.

S5: Classification of Modalities

Joëlle COUTAZ, LGI: Languages & Devices

Pipe-line model. Two levels of abstraction: physical actions & conceptual units. At the boundary between user & system, languages & devices. Physical device as modality, language as structure. For example, temperature change -> two possible conceptual units, new temperature or temperature change. Leads to two different system actions perceivable by the user.
CARE (Complementarity, Assignment, Redundancy, Equivalence) properties:
Device-equivalence: if the information can be expressed using either of the two devices. e.g., in Matis: keyboard or microphone to enter a typed or spoken request.
Equivalence is not true for the user (Jon May) -> not equivalent over (all) natural language -> equivalence is not total. (Fabio Paterno': D-equivalence ≠ task-equivalence)
Assignment: no choice to express something; e.g., in Matis, resizing a window
Redundancy: Device-equivalence + simultaneous, e.g., say 'close window' and at the same time click the window's close box
Complementarity: e.g., in Munix the user speaks the command name and types the file name to express a command.
Synergistic/Alternate/Concurrent/Exclusive Classification for the use of languages/devices
Synergistic: combination + simultaneous
Alternate: combination + non-simultaneous (either system or user constraint)
Concurrent: no combination + simultaneous
Exclusive: no combination + non-simultaneous

Q&As:
addition to MSM? another way of saying the same thing, more emphasis on languages & devices as points of contact between the user and the system.

Michaël MAY, CCI: Taxonomy of output modalities

Generalisation of the visualisation problem addressed by David Duce (see Session 4).
Modality as a concept to classify representations and not systems (a representation is that part of the presentation that conveys the information that the system intends the user to perceive and interpret).
Why a taxonomy of representations? It helps with the analysis of the information mapping problem (i.e., mapping of information types onto representation types)
Basic intuition: this is not a new problem (i.e., it has not much to do with multimodal and multimedia systems).
Useful: more informative, focussed, easy to understand, etc.
Example (the Rhode Island map): first design combines a static graphical graph, a static graphical map to indicate the density of population (using colour) and the median family income (using height of barcharts).
An alternative design uses colour and height in a redundant way to show the density of population and the income is shown by icons: it introduces a separator modality (broken line) to show the lowest incomes because the height channel is already taken up. The second design uses five modalities instead of three for the first one. Most people preferred the second one (easier to remember, more pleasing, and has redundancy).
General approach: combination of modality theory and information mapping. Identify I/O modalities leads to possible modalities and then to optimal modalities (using heuristics based on semantic properties of the representation and cognitive constraints of the user).
Analytical approach: make abstraction from individual differences; different domains of knowledge and tasks lead to generic properties across domains.
Analytical decomposition of interfaces into more simple objects to discover properties of these atomic objects
Decomposition of atomic modalities (chemistry metaphor)
the basic unit is the *unimodal atom*: meaningful representation that cannot be further decomposed. Unimodal atoms can be combined. A unimodal atom is characterised by its binary features, its specific media of expression, and its specific type of interpretation (i.e., intended interpretation). These characteristics are part of the unique profile of a unimodal atom
Binary features can be:

- linguistic/non-linguistic,
- analog/not analog—analogy is the similarity with the represented object—,
- arbitrary/not arbitrary—arbitrary means: not dependent of an already existing system of expressions—,
- static/dynamic—static means: does not violate the free inspection time of the user.

Media of expression can be: *gra*-phical, *aco*-ustic, *hap*-tic.
Type of interpretation can be: Image-*Img*, *Map*, Compositional diagram-*Comp*, Graph-*Grph*, Conceptual diagram-*Conc*, Symbolic structure-*Sym.*

Levels of inheritance in the specification of unimodal modalities: from super-level to generic level to atomic level. Mechanism describes possible combinations of features/attributes into a profile.
Advantage of the approach: systematic, tells about the properties of combinations of modalities. First analysis at the atomic level and then reconstruction by combination. Allows to predict properties of the interface (combinations of modalities).
Two-way approach: both analytical (from an existing artifact, decompose it to analyze the properties) and reconstructive (starting from atoms to build new representations). For example, a window structure (just the frame) is unimodal; with a QuickTime controller in it, it's multimodal.
Q&As:
• Link with LGI: output interaction language = CCI's representation;
• about the map example: It is not clear why people preferred the map with redundancy (but maybe it was not because of the redundancy?)

Steven VERJEANS, Leuven: InfoMap (Information mapping method & PaTerm case study)

Five-step process (possibly with backtracks):

- requirements
- refinement e.g., HTA
- Information representation (DSD)
- Information mapping: modalities + combination + interface layout
- Trade-off(s)

Case study: PaTerm is a tool for editing the dictionary of PaTrans (english to danish automatic translation system)
Step 1 (requirements)
Tasks: only the representative tasks which will then be detailed
The translation process:

- input: text
- pre-editing (just plain text)
- -> queue (because PaTrans is slow)
- translation by PaTrans
- -> queue (post-editing even slower- done by a human)
- post-editing (proof checking)
- output: 'authorized' translations

Remote study of an interface/reconstructive (worst case!)
Step2 (detailed Task Analysis)
Translation system dictionary adapted to the context; automatically built from a corpus
Three situations for modification of the dictionary: building a new dictionary, adding new terms in the pre-editing phase, adding new items in the post-editing phase
Step-by-step task specification: describes the task details and the parameters of the task (pre-task organisation)
Step 3 (Information Representation)
DSD is used to represent the task analysis (helps clarifying)
Step 4 (Information mapping)
Choose modalities to represent the information. Choose also the combinations of these modalities and the layout of the interface.

relies on modality theory : from representational features to modality attributes
+ requirements on the display
other requirements are involved to constrain use/structure of information channels
Step 5 (trade-off alternatives)
e.g., criteria for choice between options
layout consistency, semantic consistency, input consistency, recognisability, flexibility
other example: criteria for buttons (click or touch)
button size, button grouping, button distance
Use of QOC to elicit possible choices and help make the choice
Q&As:
• (in an example that was presented) preferred option was (c): was not chosen by users but just people seeing the layout at a glance. It could be that since all four elements look the same, the layout esthetically looks better.
• no symmetry between output/input modalities; input modalities not analysed yet (so -> kbd, mouse, etc)
• interesting point is that it’s explicitly task-related