Reliable Visual Analytics within a Verification and Validation Management

Wolfram Luther, Ekaterina Auer

Electrical Engineering and Computer Science
University of Applied Science Wismar
Motivation

This talk aims at widening the focus from one of reliable numerical algorithms and efficient implementations using exact real arithmetic to a broad system modeling approach and evaluation design.

A recently described four tier verification and validation management defines requirements for categorization and classification of processes as a result of precise assessment procedures and addresses recommending techniques, user interaction and collaboration via adequate human machine interfaces.

Huge data and program code require new visual analysis methods. Reliable visual analytics is paired with an assessment of (meta)data and code quality, adequate data types and methods to propagate and bound uncertainty.
Overview

- Motivation and Application Context
- Visual Analytics
- Visualization of Uncertainty - a Taxonomy
- Ontology-based Architecture
- A Tool Chain – Realized Software
- Use Cases
  - Femur Prosthesis (PROREOP) – Computational model verification
  - GPS Sensing and Localization, Spatial Decision Making using DST
  - Risk Communication and Perception – House of Risk
  - Reliable Analysis of Steel Samples SILENOS©
  - Modeling and Implementation of a Microscopic Traffic Simulation System using MAUDE
  - Virtual Museums and Labs – ViMEDEAS
Motivation and Application Context

- Real world process modeling and simulation
- Application of V&V assessment to improve reliability
- Visual Analytics as a new interesting aid
  - Ensemble data, independent process runs
    → Various analysis goals
  - In most cases, (result) verification is not an issue
  - Collecting, processing, quantifying, displaying uncertain data
  - Various modeling approaches
- Adapted interfaces and query techniques
- Toolchain for the whole software support
- Evaluation scope focusing on users, models, and data
Validation & Verification Assessment

Methodologies
Tools
Standardization
Accuracy/Reliability classification

Design

Real World

Process/System Task

Analysis

IVP/PDE

Formal Model

Verification

Simulation

Visual Analytics

HCI

Computer-based Model

Implementation

prototype program

Computationalists

Define verification layers
code verification/software quality/testing
performance analysis
accuracy assessment
analytical solution
tools with result verification
sensitivity analysis
a priori / a posteriori error bounds
benchmark examples
independent computations

Experimentalists

Define validation layers
validation metrics
experiments simulations
statistical approach
 calibration
performance/safety

Validation

System Modeling, Simulation, and Verification Cycle
Processes, Systems, Tasks

Modelists

Define application domain
relevant parameters and ranges
adequate data types
quality of input data
types of uncertainty
uncertainty management methods
time scales
inherent logic refinement
Visual Analytics

• Complex processes use/produce huge, heterogeneous data
• Variables and parameters exhibit aleatoric and epistemic uncertainty
• Important issues are user-controlled selection of data types, problem solving approaches, design of human-machine interaction
• Hierarchical ensemble data are
  • Used in process model descriptions
  • Processed by sophisticated software tools
  • Displayed using repeated multiple views or incremental approximation
  • Analyzed in various ways to rate the process outcome
• Various evaluation approaches
  • Rule-based, role-based or knowledge-based
  • Usability: interaction style, task model, data handling, perception
  • Verification and validation taxonomy for computerized system models
• Taxonomy wrt. data dimension – domain, range of $f$
  • 1D (for scalars)
  • 2D, 3D (for spatial vectors)
  • ND (for non-spatial, multivariate, and time-varying data)
• Uncertainty added as a geometric form (line, bar chart, thick surface)
• Also: color maps, glyphs or isosurfaces (level sets) (unit of space - arrow, unit of color, number of graphical elements)
• More complex functional process descriptions are not included (Potter et al. 2012)
• Missing user interaction, requirements concerning hardware architecture or context detection, perception issues
Brodlie (1992, 2012) introduced E (exact), U (uncertain) notation, which also records the number of independent variables or parameters as a subscript and uses a superscript to indicate the type of dependent variables.

Examples

- $E_1^S$: scalar function of one variable
- $E_1^{kS}$: $k$ scalar functions
- $kE^S$: multi-field scalar data ($E_{\geq 0}^{kS}$)
- $E^V$: vector data.
Examples

- Function \((x,y; f,m)\), 2D earth map; seismic loss event frequency (color) and loss magnitude (size)
- Uncertainty contour band indicating the boundaries of the 95% confidence interval
- Diagram relation \((x,y; z)\): \(E^z_{xy}\) with integer or alphanumeric data \(x,y,z\) or intervals \((x,y; v)\): \(U^v_{xy}\), \(x\) damage, \(y\) frequency, \(v\) acceptance (color)
- Uncertainty isosurface with color \(U^v_{xyz}\)
- Animation of diagrams over time \(E^v_{xyt}\)
- A set of relations (edges) between nodes visualized as a graph structure \(S'\).

Isosurface: Brodlie et al 2012 – Figure 8
More general: Ordered couples of input-output data \((d_{i1}, ..., d_{ik})\), \((d_{o1}, ..., d_{oj})\) and metadata (descriptors) \((m_{i1}, ..., m_{ir})\), \((m_{o1}, ..., m_{os})\) of the underlying relation (set) \(R\), process \(P\) or structure \(S'\). Selecting a few \(d_j\) or a projection restricting certain variables or parameters to a bounded interval, a set or a precise value of a variable (parameter) reduces dimension Uncertainty \(U^V_x\) could be represented as an interval or error bar \(z_x\) over a discrete set \(X\), a truncated upper and lower PDF or interval mean and the standard deviation of a PDF. Varying contour color or thickness and surface opacity illustrate regions of uncertainty across the spatial domain, often augmented with uncertainty annotations.
• Windows, icons, menus, pointing (WIMP) interfaces utilizing mouse and keyboard-based interaction on screens are well suited for presenting 2D content
• Post-WIMP interfaces allow for new interaction paradigms for the navigation and manipulation of 3D Virtual Reality Environments and visualizations.
• They use 3D devices to navigate and select objects and to grab or grasp and manipulate items, e.g., an elastic arm and a virtual hand; users move around items, detect interesting viewpoints or areas of interest.
• Extensions of the W3C task meta-model are needed to face new interaction styles, which do not depend on classical 2D UI elements. The meta-model offers a hierarchical structure among tasks and provides several operators to define temporal relationships between tasks.

• Extended post-WIMP task models (PWTM) have to include adequacy of interaction elements, flexibility in partitioning the task among multiple actors, multimodal fault-tolerance and error-avoiding dialogues with forward and backward error recovery to cover uncertainty issues.

• PWTM should have profiles depending on the application type, for example, cooperation or virtual reality.
Immersive Environments

• Vivid representation of objects, particularly real life scenarios, such as people walking in streets: traffic, landscapes, sport simulations, power plants or other technical processes
• Navigation, selection, surrounding and manipulation of items in a 3D environment
• Situational awareness and new forms of collaboration in problem management and decision making—shared workspaces, multi touch-tables, and co-creation
• Environment models seasons, day and night, weather, and various landmarks or navigation aids.
Ontology-based Architecture

Workflow

Gathering Data and Metadata

Creating a Model

Queryable Ontology

Visualizer

Output devices
Application Software

• Steel application → analyzes data collected about non-metallic inclusions and other defects in steel samples: Image processing, Particle Detecting and Analysis System, Inclusion Processing Framework Viewer IPF 2.0 (Chr. Buck, M. Thurau, 2011-2016)
• House of Risk → devoted to individual threats, thematically classified and placed in an indoor or outdoor context. It will also address public threats and macro-catastrophes etc. (A. & L. 2016 - )
• GIS-applications → Uncertainty modeling DST, traffic, localization, network planning (G. Rebner, B. Weyers, J. Frez, 2012 - )
• Femur prosthesis surgery → Data grabbing, reliable superquadrics modeling, visualization – integrated framework for verified geometric computations (R. Cuypers, St. Kiel, 2009 – 2014)
• Microscopic traffic modeling and simulation system → Code verification, model validation MAUDE (J. Brügmann, 2013-2015)
Steel Process Data

- Steel Metadata: Various steel grades
- Process Parameters: Intentional settings or measurements taken during monitoring
- Sample Parameters: Parameters for the milling machine that slices the steel surface, and more complex statistical descriptors of the defects found, such as the sample cleanliness
- Defect Parameters, descriptors and volume data for each defect
- Incremental Approximation Analysis
- Result Reports depending on users’ interest
- Trend and Sensitivity Analysis: How the defect data (positions, sizes, types, number) change when process parameters are changed?
- Variability and Uncertainty Analysis
Defects in steel samples with the IPF Viewer

- Analyze the ensemble data (geometrical data, descriptors etc.) in various ways
- Identify samples and defects that differ from others by position, size, type and number
- Find trends about how different process parameters influence the steel samples
- Perform variance analysis to examine natural fluctuations within the samples and desired variations

Context data about the steel and sample

Pie chart of the containing defect types

Aggregation statistics about contained defects

Scatter plot showing positions of defects, colored by defect type

Histograms about different defect dimensions

Data mining results: most interesting defects of the sample
Multidimensional grouping comparable to OLAP cubes in real time. In addition to numerical analysis of the aggregates, the IPFViewer visualizes the data cells with repeated multiple views (flat on the screen) and presents the data mining results for each group.
Various models: *physical reality, mental model, formal model, computational model (real world process or task)*

- Conceptual design
- Conceptual and perceptual model
- Interaction and action design
- Interaction and *action logic*
- Interface design
- Observation, exploration, interaction, de(re-)construction
- User model
- Interface, *tutor and expert module*

*Ex: Visual analysis of finite geometry and raster algorithms*
Virtual Museums (Sacher, Biella, Otten, Weyers 2015/16)

- **ViMEDEAS Virtual Museum Exhibition Designer Using Enhanced ARCO Standard**
- **ViMCOX Virtual Museum and Cultural Object Exchange Format**

- Generative approach facilitates the implementation and combination of tailored authoring tools (2D/3D/WYSIWYG-Generators/Services)
  - Floor Planner
  - Museum Layout Designer
  - Exhibition Designer
  - Interaction/Collaboration Designer
  - Mobile Data Acquisition
  - Middleware (Content Connector, Metadata Mapper, Publisher, Generator)
Showcase – Mahn- und Gedenkstätte Düsseldorf

- The Virtual Leopold Fleischhacker Museum
- Usability and Utility Evaluation (Lab), ISO/IEC 9126
- Kiosk-system Exhibition in Düsseldorf
- Cave Automated Virtual Environment Exhibition
- Comparative Evaluation (On-site)
- Visual analysis of gravestones
- Currently we are working on an Armenian Khatchkar museum
GIS - Localization Approaches

- Initial object position and direction \((x; y; \theta)\)
- Autonomous robot: vision systems, beacons, landmarks, or GPS
- Markovian Localization Framework
  - (Extended) Kalman or Bayesian filter techniques
- Wide Area Augmentation System (WAAS), Differential GPS (DGPS), and Assisted GPS (AGPS)
- Inertial Navigation Systems (INS)
- Combining guaranteed interval methods with random approaches
Possible Basic Ontology for GIS Applications

- **Thing**
  - **GPS-satellites**
    - count
    - local correction
    - XDOP
  - **Physical space**
    - has environment
    - defined by Dimension
  - **DS_concept**
    - contains object(s)
    - Basic probability assignment
      - **DS_source** (xsd: string)
      - **DS_belief** (xsd: interval)
      - **DS_plausibility** (xsd: interval)
      - **DS_mass** (xsd: interval)
  - **Dimension**
    - 2: horizontal space H,x,y
    - 1: vertical space V,h
    - 1: temporal space T,t
    - 3: positional space P,x,y,h
    - 4: geometric space G,x,y,h,t

26.11.-01.12.2017

Dagstuhl 2017
Typical Queries

- Compute *Belief* and/or *Plausibility* that a horizontal/vertical/temporal/positional/geometric room contains an object.
- Inverse query: find a space where Belief and/or Plausibility is a greater/smaller part of a certain interval.
- Compute a Belief and/or Plausibility for more than one object.
GeoSPARQL – Basic Classes

• **Spatial Object**

```turtle
geo:SpatialObject a rdfs:class, owl:class;
  rdfs:isDefinedBy <http://www.opengis.net/spec/geosparql/1.0>;
  rdfs:label "Spatial Object"@en;
  rdfs:comment "This class represents everything that can have a spatial representation. It is super class of Feature and Geometry."@en .
```

• **Feature**

```turtle
geo:Feature a rdfs:class, owl:class;
  rdfs:isDefinedBy <http://www.opengis.net/spec/geosparql/1.0>;
  rdfs:label "label"@en;
  rdfs:subClassOf geo:SpatialObject;
  rdfs:disjointWith geo:Geometry;
  rdfs:comment "This class represents the top level feature type. This class is equivalent to GFI_Feature defined in ISO 19156, and it is super class of all feature types."@en .
```

• **Geometry**

• Are all Spatial Objects that are no features (see above)
• GeoSPARQL further specifies a Geometry Class Hierarchy
GeoSPARQL specifies three sets of topological relation types based on a simple (GeoSPARQL own) definition, Region connected calculus (RCC8), and Dimensionally Extended nine-Intersection Model (DE-9IM, Egenhofer), but is not restricted to those.

<table>
<thead>
<tr>
<th>Relation Name</th>
<th>Relation URI</th>
<th>Domain/Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>equals</td>
<td>geo:sfEquals</td>
<td>geo:SpatialObject</td>
</tr>
<tr>
<td>disjoint</td>
<td>geo:sfDisjoint</td>
<td>geo:SpatialObject</td>
</tr>
<tr>
<td>intersects</td>
<td>geo:sfIntersects</td>
<td>geo:SpatialObject</td>
</tr>
<tr>
<td>touches</td>
<td>geo:sfTouches</td>
<td>geo:SpatialObject</td>
</tr>
<tr>
<td>within</td>
<td>geo:sfWithin</td>
<td>geo:SpatialObject</td>
</tr>
<tr>
<td>contains</td>
<td>geo:sfContains</td>
<td>geo:SpatialObject</td>
</tr>
<tr>
<td>overlaps</td>
<td>geo:sfOverlaps</td>
<td>geo:SpatialObject</td>
</tr>
<tr>
<td>crosses</td>
<td>geo:sfCrosses</td>
<td>geo:SpatialObject</td>
</tr>
</tbody>
</table>
Query Functions Based on Topology

- Some Boolean query functions defined for the *Simple Features* relation family
- Multi-row intersection patterns should be interpreted as a logical OR of each row
- Each function accepts two arguments (geom1 and geom2) of the geometry literal serialization type specified by *serialization* and *version*.
- Each function returns an *xsd:boolean* value of **true** if the specified relation exists between geom1 and geom2 and returns **false** otherwise

<table>
<thead>
<tr>
<th>Query Function</th>
</tr>
</thead>
</table>

OGC GeoSPARQL - A Geographic Query Language for RDF Data
Graphical representation of the solution set describing possible locations and orientations of a truck.
Extended GeoSPARQL for Uncertain GIS

Data Graph

Query

SPARQL Engine

```

```

Dempster Shafer Toolbox
Truck Example

Different rendering techniques and visualization approaches to the (convex) hull geometries
Outlook: House of Risk

**Goal:** Process, format and present relevant everyday threats from several risk classes that appear in specific contexts with typical parameters in a virtual house of risk.

**Tasks of the presentation system:** Deal with the location or time dependency of threats and discriminate among types of risks using a variety of threat categories: technological, geopolitical, societal, economic/financial, environmental, and threats to the life and health of private citizens.

**Planned realization:** A layered architecture approach providing risk data and metadata in a virtual reality environment allowing for appropriate 2D/3D visualization/interaction styles and for easy extension with new risk/threat concepts in a predefined format.
Workflow for analyzing and visualizing risk models and environmental entities under uncertainty
Visual analysis of bird swarms as a risk for aviation
Total Hip Arthroplasty Simulation (R. Cuypers, 2011)

Complete V&V assessment – Dagstuhl 2008
• OLSIM v 4.0 (Bigloo)
• The Maude system executed the simulation model to reproduce the well known fundamental diagrams of the NaSch microscopic traffic model (velocity/flow to density with(out) collision)
• Code verification with Maude
• Collision avoidance
• Travel time validation
• Performance validation
• Some deficiencies in the number generator
V&V management was applied within a workflow for designing, modeling, and implementing various processes. Toolboxes for the spatial decision making, steel artifacts, femur prosthesis, virtual labs, traffic simulation etc. were implemented for supporting reliable visual analytics. Evaluation, validation considered from various vantage points concerning design, code/numerical result verification, software testing, recommendation/usability were conducted. Numerical result verification is only one issue, bad/huge data, user interaction, task and process modeling and visual analytics should also be taken into account.
• B. Weyers, W. Luther, Risk Communication and Perception in Low- and High-Immersion Virtual Environments, APSSRA6, 28-30 May 2016, Shanghai, China, H.W. Huang, J. Li, J. Zhang & J.B. Chen (editors) (2016)
• G. Rebner, D. Sacher, B. Weyers, W. Luther, Verified stochastic methods in geographic information system applications with uncertainty, Structural Safety 52 (2015) 244-259
• K. Potter, P. Rosen, and C. R. Johnson, From Quantification to Visualization: A taxonomy of uncertainty visualization approaches, IFIP Advances in Information and Communication Technology Vol. 377 (2012) 226-249
• GeoSPARQL – A Geographic Query Language for RDF Data, http://www.opengis.net/doc/IS/geosparql/1.0
• Publications and projects SCG: http://www.scg.inf.uni-due.de/
Thank you for your interest!

Thanks to
Dr. J. Brügmann, Dr. Chr. Buck, Dr. D. Biella, J. Frez, Dr. D. Sacher, Dr. St. Kiel, M. Thurau, Dr. B. Weyers for their contributions
our partners Prof. Beer (U Hannover), Prof. Aschemann, Dr. Rauh (U Rostock), Profs. Baloian, Ochoa, Pino (U Chile)
and to the EU Ziel 2 and DAAD, DFG for funding our work.
Abstract

Reliable visual analytics within a verification and validation management

In recent years, we have addressed together with international and industrial partners a Verification & Validation Assessment with special emphasis on system and process modeling under uncertainty, code verification, numerical accuracy and performance testing including also user-centered recommendation technics or reliable visual analytics.

In this talk, we highlight an ontology-based architecture with a query engine and modern human machine interaction for various use cases: GPS sensing and localization, spatial decision making, risk communication and perception, analysis of steel samples using SILENOS®, virtual museums and labs software ViMEDEAS, Biomechanics: femur prosthesis (PROREOP) as well as modeling and implementation of a microscopic traffic simulation (OLSIMv4).