Patterns as Formulas: Patterns in the Digital Humanities

Johanna Barzen, Frank Leymann

Institute of Architecture of Application Systems,
University of Stuttgart, Germany
{barzen, leymann}@iaas.uni-stuttgart.de

BibTeX

```
@inproceedings {INPROC-2017-17,
    author = {Johanna Barzen and Frank Leymann},
    title = {{Patterns as Formulas: Patterns in the Digital Humanities}},
    booktitle = {Proceedings of the Ninth International Conferences on Pervasive
Patterns and Applications (PATTERNS)},
    address = {Athen},
    publisher = {Xpert Publishing Services},
    pages = {17--21},
    isbn = {978-1-61208-534-0},

    url = {http://www2.informatik.uni-stuttgart.de/cgi-
bin/NCSTRL/NCSTRL_view.pl?id=INPROC-2017-17&engl=}
 }
```

The full version of this publication has been presented at PATTERNS 2017.
http://www.iaria.org/conferences2017/PATTERNS17.html

Patterns as Formulas:

Patterns in the Digital Humanities

Johanna Barzen and Frank Leymann
Institute of Architecture of Application Systems
University of Stuttgart
Stuttgart, Germany
e-mail: {barzen, leymann}@iaas.uni-stuttgart.de

Abstract—**During the last years, in particular due to the Digital Humanities, empirical processes, data capturing or data analysis got more and more popular as part of humanities research. In this paper, we want to show that even the complete scientific method of natural science can be applied in the humanities. By applying the scientific method to the humanities, certain kinds of problems can be solved in a confirmable and replicable manner. In particular, we will argue that patterns may be perceived as the analogon to formulas in natural science. This may provide a new way of representing solution-oriented knowledge in the humanities.**

Keywords-pattern; pattern languages; digital humanities; formalisation.

I. INTRODUCTION

A fundamental aspect of the scientific method (i.e., the method of the natural sciences) is repeatability. Repeatability allows to gain two key goals of research: objectivity and solvability.

Repeatability is the basis for verifiability of research results. Verifiability allows to establish objectivity in the sense of not having to rely on trusted authorities (i.e., well-accepted domain experts) expressing their subjective insights as research results. As a consequence, everybody can re-enact and track the way a research result has been achieved.

Often, a research result itself that has been obtained by applying the scientific method has an aspect of repeatability too. A corresponding result is represented as a procedure to solve recurring problems of a certain kind. Such a result is often expressed as a formula, and the solvability of the problem is achieved by applying this formula.

Also, the humanities, in particular the Digital Humanities [1], have domains in which objectivity and solvability (in the sense stated above) are important goals. As shown in the following, the scientific method may be applied in such domains to achieve both, verifiability of results as well as results that have a "solution character".

First, in Section II, we give evidence that the concept of patterns is a proper vehicle to establish the solvability facet of research results in the humanities. Second, in Section III, we show that applying the scientific method in data intensive domains of the humanities establishes the objectivity facet of research results. Section IV concludes the paper.

II. PATTERN AS FORMULAS

Solving problems in natural sciences by means of formulas is in close analogy to using patterns in other domains like architecture [2] or software engineering [3][4], for example. Applying a formula means to follow a certain proceeding: identifying the intent to solve a specific problem, determining the solution sketch and apply it to the actual context of the problem to be solved to result in the concrete solution of the problem.

A. Example: How to Solve Quadratic Equations

Assume that one has the **intent** to solve the problem of determining the roots of the following quadratic equation (1):

$$x^2 + 5x + 4 = 0 \tag{1}$$

What is done first is to consult a formulary to find the **sketch** of how to solve arbitrary quadratic equations $ax^2 + bx + c = 0$ and the quadratic formula (2)

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{2}$$

for solving (1) is found, i.e., for determining the roots x_1 and x_2.

Next, the ***context*** of the problem to solve the specific equation $x^2+5x+4=0$ needs to be determined, i.e., the actual coefficients (3) of the concrete equation have to be determined:

$$a=1,\ b=5,\ c=4 \tag{3}$$

After understanding the sketch and the context, the corresponding quadratic formula (2) and the actual coefficients (3) have to be combined, i.e., the coefficients *a, b, c* in the quadratic formula (1) have to be substituted by the actual values $a=1$, $b=5$, and $c=4$. Together, the ***solution*** description (4) of the problem has been determined:

$$x_{1,2} = \frac{-5 \pm \sqrt{5^2 - 4\cdot 1\cdot 4}}{2\cdot 1} \tag{4}$$

This results in the ***concrete solution*** of the original problem, i.e., the roots of the quadratic equation (1), see (5):

$$x_1=-1,\ x_2=-4 \tag{5}$$

B. *Using a Formula Means Applying a Pattern*

This proceeding of determining the roots of a quadratic equation can be described by a document that follows the template of a pattern (as introduced in [2][3][4]): First, a pattern has a name that uniquely identifies the **problem** to be solved. Second, it specifies the **intent** of solving the particular problem (e.g. finding the roots of a quadratic equation). Then, it describes in a **sketch** how to solve the problem (i.e., the quadratic formula). Next, it lays out how to determine the **context** in which the solution can be applied. Finally, the **solution** works out how the before-mentioned information is put together to solve the problem. Figure 1 depicts these elements of the corresponding pattern document (where the context-section and the solution-section are combined). Note, that these ingredients of the pattern document make use of only the essentials of a pattern template: pattern languages typically capture more information [2][3][4].

C. *Example: Solving a Costume Design Problem*

Patterns are used in different domains in the humanities. But the term "pattern" is often rather problematic because of the different meanings it refers to [5]. But used in the sense introduced by C. Alexander [2], patterns are a convincing tool to capture knowledge and make this knowledge easily accessible. In the MUSE (MUSE - MUster Suchen und Erkennen, engl.: pattern search and identification) project [6], patterns are used to document solutions of costume design problems in films.

This project is about solving a problem from the humanities (more specifically from the media studies), namely proving the existence of a costume language and providing such a costume language for several genres. The individual costumes found are documented as patterns (see Figure 2); note, that the pattern content shown is just an example and not yet a verified pattern.

As before, the pattern document begins with the name of the problem. It describes the intent of solving the problem. The sketch presents the essentials of the solution, and the context describes the circumstances of applying the solution. Finally, the solution discusses in details how the costume is built – in this case, a figure depicts all the primitives of the costume and the order in which they are worn.

By following the pattern, a solution to a certain costume design problem is constructed: just like following the pattern for solving quadratic equations. Thus, using patterns in the humanities to express research results brings the power of applying formulas to the humanities, i.e., it establishes the solvability facet of research result in the humanities.

III. The Scientific Method

In a nutshell (and admitting, this is a very simplified view), the scientific method consists of the following steps [7][8]: observation, data capture, data analysis, and formalization (or abstraction).

An *observation* can be based on planned experiments, systematically watching phenomena in nature etc. Often, observations are caused by a hypothesis resulting from theoretical reasoning. *Data capturing* refers to the stringent logging of information resulting from observations. *Data analysis* takes a close look to the captured data in order to find regularities. By means of *abstraction* or *formalization* found regularities are finally expressed as laws or formulas.

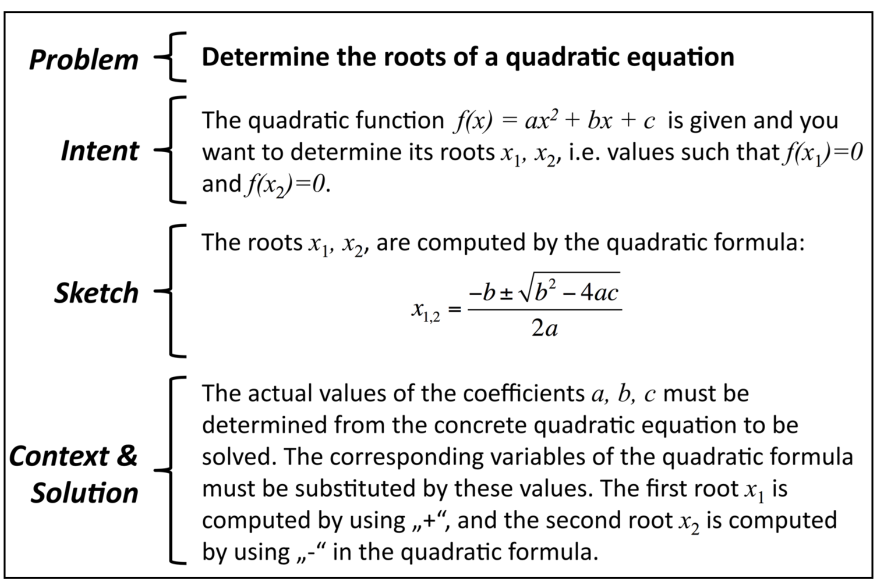

Figure 1. Using a Formula as Applying a Pattern

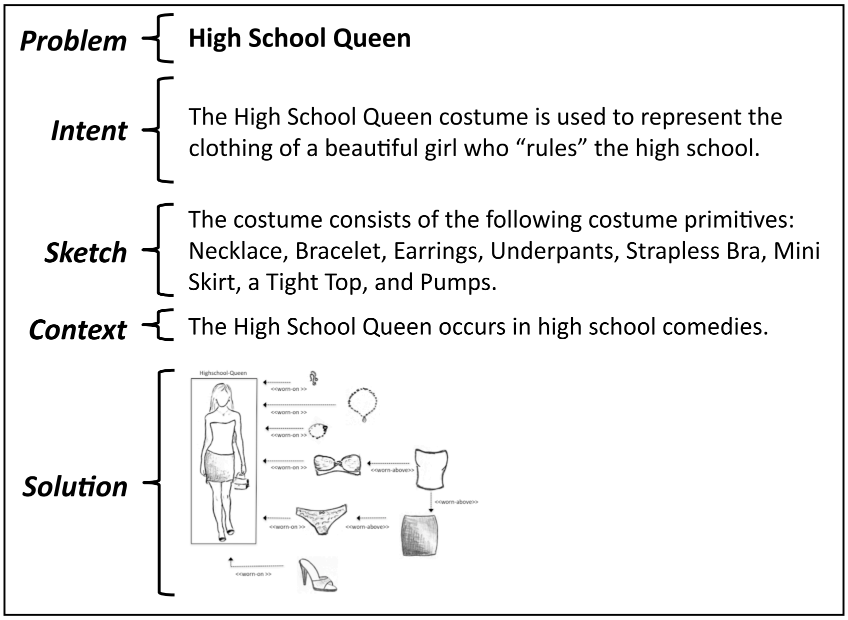

Figure 2. A Costume Described by a Pattern

A. Example: Determining Planetary Motion

Historically, Tycho Brahe observed celestial positions of planets (especially Mars) and captured the corresponding data in logs [9]. Johannes Kepler analyzed this data [9] and inferred Kepler's Laws of planetary motions [10]. For example, the first Kepler Law is: "The orbit of each planet is an ellipse with the sun in one of its foci." Kepler also formalized his laws, i.e., he abstracted his laws as mathematical formulas. The first Kepler Law, for example, is formalized as (6)

$$r = \frac{p}{1 + \varepsilon \cos\theta} \tag{6}$$

Thus, as Figure 3 depicts, the discovery of Kepler's Laws of planetary motions follows the proceeding of the scientific method sketched above: observation, data capture, data analysis, and formalization (or abstraction).

B. Example: Determining Costume Languages

The problem of determining costume languages has been described in [11]. In [12], a method (called the MUSE Method) has been described for deriving costume languages. This method is a refinement of the (simplified) scientific method sketched in Figure 3: The left-hand side of Figure 4 shows the MUSE method as described in [12] (ignoring aspects not relevant in the context of this paper), the right-hand side shows the scientific method, and the correspondence between steps in the MUSE method and steps in the scientific method are indicated as arrows.

In the MUSE method, characters of a film corpus are identified. This is done by watching the corresponding films – which means that observation is taking place. While watching the films, the cloths of the identified characters are described – which is data capturing. Based on the captured data about clothes, costumes are identified – which is data analysis. Identified costumes are then abstracted into patterns – which can be considered as formalization as argued above.

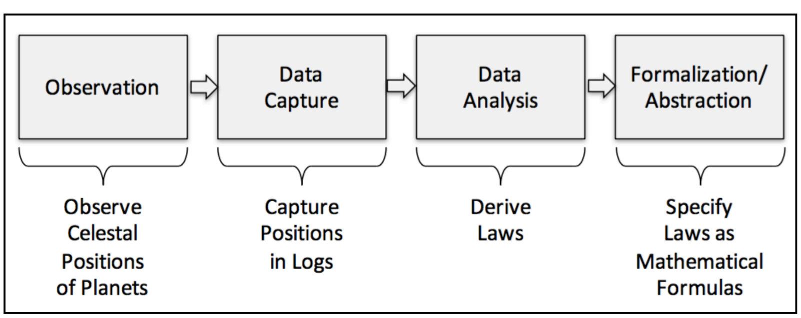

Figure 3. Principle Proceeding in Deriving Kepler's Laws

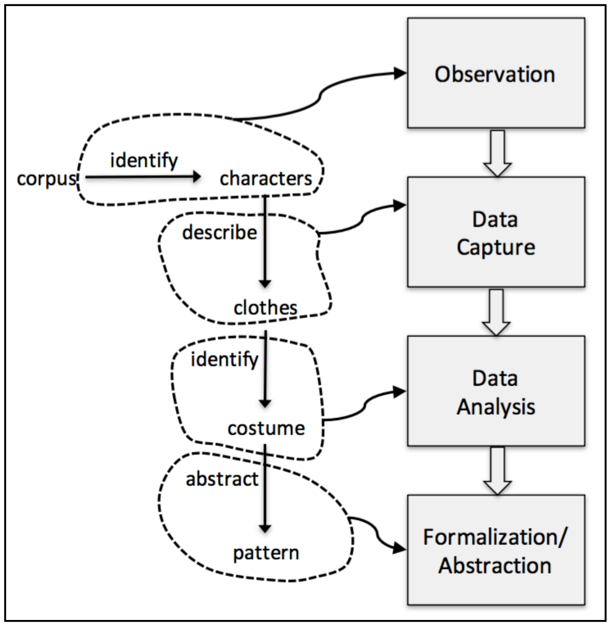

Figure 4. The MUSE Method as Scientific Method

An overall software system that supports the MUSE method has been build, and that system is applied in the domain of film studies in the humanities [13]. In particular, the discovery of costume languages is supported. This system allows to describe films, their characters and the cloths of the characters via a graphical user interface and stores this data in a database. The structure of this database as well as the domains of its central attributes are modeled by taxonomies and ontologies [12][14]. The analysis of the data is supported by means of data warehouse and OLAP (Online Analytical Processing) technologies [15], as well as by means data mining technologies [16]. The representation of the abstracted patterns and their relations (i.e., the resulting pattern language) is stored in a pattern repository [13].

The software system can be used to verify research results, thus, contributing to objectivity: Everybody can browse the captured data to assess its quality; the captured data can be analyzed over and over again to confirm the discovered regularities within the cloths of the films; the patterns reference the costumes they have been abstracted from, which support to track the abstraction of similar cloths to costume patterns. The latter, by the way, does also contribute to solvability: a pattern does not only describe abstractly the structure of a certain costume but also provides a set of concrete cloths to realize this costume.

Formalization in MUSE goes even beyond describing research results as patterns. Cloths themselves as well as their constituents are considered as words of formal languages [12]. This allows, for example, to check whether newly discovered cloths are in an already established tradition.

IV. Conclusion and Outlook

This brief contribution argued in favor of applying the scientific method in the humanities. In doing so, repeatability of research results – especially the facets of objectivity and solvability – will be emphasized. Patterns have been presented as an analogon to formulas as an integral part of the scientific method.

It has been shown how the scientific method has been applied in the film studies. The scientific method in general, and the MUSE method in particular is currently being applied in musicology [17] - in an effort called MUSE4Music.

There are important implications of applying the scientific method to domains or in ways it has not been established for, and these implications are independent from the fact whether the scientific method is applied in the humanities or in natural sciences. In order to ensure repeatability, the data as well as the algorithms used to analyze this data to achieve the results must be published [18]. This is by far not yet widely accepted because the data and algorithms are often considered proprietary or a “production secrete” for achieving research results. This is an obstruction that has to be overcome.

Note

This paper has been pre-published as a technical report of University Stuttgart no. 2016/01.

References

[1] A. Burdik, J. Drucker, P. Lunenfeld, T. Presner, and J. Schnapp, “Digital_Humanities” Cambridge/London: The MIT PRESS, 2012.

[2] C. Alexander, “The Timeless Way of Building,” New York: Oxford University Press, 1979.

[3] G. Hohpe and B. Woolf, “Enterprise Integration Patterns,” Addison-Wesley, 2004.

[4] C. Fehling, F. Leymann, R. Retter, W. Schupeckand, and P. Arbitter, “Cloud Computing Patterns,” Wien: Springer, 2014.

[5] D. Dixon, “Analysis Tool or Research Methodology: Is There An Epistemology for Patterns?” in D. Berry (ed.): Understanding Digital Humanities, Pilgrave Macmillan 2012, pp. 191-209.

[6] Available from: http://www.iaas.uni-stuttgart.de/forschung/projects/MUSE/indexE.php

[7] S. Charey, “A Beginners Guide to Scientific Method” (4th Edition) Boston: Wadsworth, 2011.

[8] T. Hey, S. Tansley, and K. Tolle (Ed.), “The Fourth Paradigm: Data-Intensive Scientific Discovery” Microsoft Research 2009.

[9] F. Hund, “Geschichte der physikalischen Begriffe (engl. History of physical terms),” Mannheim: Bibliographisches Institut, 1972.

[10] M. Schneider, “Himmelsmechanik (engl. Celestial Mechanics),” Mannheim: Bibliographisches Institut, 1981.

[11] D. Schumm, J. Barzen, F. Leymann, and L. Ellrich, “A Pattern Language for Costumes in Films” Proceedings of the 17th European Conference on Pattern Languages of Programs (EuroPLoP), pp. C4-1–C4-30, 2012. Available from: http://www.europlop.net/sites/default/files/files/proceedings/EuroPLoP2012_companion_proceedings.pdf 2017.01.28

[12] J. Barzen and F. Leymann, “Costume Languages As Pattern Languages” in Proceedings of Pursuit of Pattern Languages for Societal Change (PURPLSOC) - Preparatory Workshop, Krems: epubli GmbH, pp. 88-117, 2014.

[13] C. Fehling, J. Barzen, M. Falkenthal, and F. Leymann, “PatternPedia - Collaborative Pattern Identification and Authoring” in Proceedings of Pursuit of Pattern Languages for Societal Change (PURPLSOC) - Preparatory Workshop, Krems: epubli GmbH, pp. 252-305, 2014.

[14] J. Barzen, “Taxonomien kostümrelevanter Parameter: Annäherung an eine Ontologisierung der Domäne des Filmkostüms (engl. Taxonomies of costume relevant parameters: approaching an ontology of the domain of the film costume),” University Stuttgart, Technical Report 2013/04, 2013. Available from: ftp://ftp.informatik.uni-stuttgart.de/pub/library/ncstrl.ustuttgart_fi/TR-2013-04/TR-2013-04.pdf 2017.01.28

[15] M. Falkenthal et al, “Datenanalyse in den Digital Humanities – Eine Annäherung an Kostümmuster mittels OLAP Cubes (engl. Data analysis in Digital Humanities - an approach to costume patterns using OLAP Cubes),” in Proceedings Datenbanksysteme für Business, Technologie und Web (BTW 2015), Bonn: Lecture Notes in Informatics, pp.663-666, 2015.

[16] M. Falkenthal et al, “Pattern Research in the Digital Humanities: How Data Mining Techniques Support the Identification of Costume Patterns” Computer Science - Research and Development, Vol. 22 (74), Heidelberg: Springer, 2016, DOI 10.1007/s00450-016-0331-6.

[17] J. Barzen et al., “The vision for MUSE4Music. Applying the MUSE method in musicology“, Computer Science - Research and Development, Vol. 22 (74), Heidelberg: Springer, 2016, DOI 10.1007/s00450-016-0336-1.

[18] F. Leymann, “Linked Compute Units and Linked Experiments: Using Topology and Orchestration Technology for Flexible Support of Scientific Applications” Software Service and Application Engineering – LNCS7365, Springer, pp. 71-80, 2012.