

Strawberries are Nuts

Tim Frey
OvG University Magdeburg,
Germany
tim.frey@tim-frey.com

Marius Gelhausen
TU Darmstadt, Germany
marius.gelhausen@gmx.de

ABSTRACT

Separation of concerns is a central element for program comprehension. This note briefly explains why human categorization can be interesting for program comprehension.

Categories and Subject Descriptors

D.2.3 [Software Engineering]: Coding Tools and Techniques;
D.2.7 [Software Engineering]: Distribution, Maintenance, and Enhancement

General Terms

Human Factors

Keywords

Separation of Concerns, Categorization

1. INTRODUCTION

Strawberries are not berries at all; the small seeds are nuts. Therefore strawberries belong to the category of nuts, but even in the awareness of this fact most people would not count a strawberry as a typical example of a nut. This paper addresses the problem that, category concepts in human categorization are not as sharp and absolute as it intuitively seems. The contribution of this note is to propose further investigation of the area of categorization in the context of separation of concerns (SOC).

2. CATEGORIZATION SPACE

Human categorization is the process of grouping objects together into categories, based on similarities. It seems a category is based on a kind of family resemblance. This means, not all objects in a category share a certain feature set; they are more like a family. For example a child has features of a mother and a father. Even though, the mother has nothing in common with the father, they belong all together to the category of a specific family [7]. One of the main applications of categorization is that membership of objects in a category, enables the possibility to use prior knowledge about a category for new objects that belong to the same [3]. Considering the strawberry - nut sample, categorization is not perfect. Even knowing the facts, category associations seem in this case fuzzy and likely wrong. Also experiments showed, that additionally a suboptimality in both categorization and identification exists [2]. Likely, a human would count a strawberry, even in knowledge of the facts, to different categories at the same time: Nuts and berries. Both categories are used to build inferences about the features [8] like taste or reproduction.

3. CONCERNS ARE CATEGORIES

A widespread idea to advance programming is the paradigm of SOC [4]. A concern is understood, to be any matter of interest in a software system [6]. When applying the idea of SOC a

developer tries to code concerns into separate modules. Figure 1 shows the mapping of the white concern space to the grey solution space. In *a* the mapping is done perfectly. In *b* the problem is shown, that different concerns are coded into one module. Finally *c* shows that one concern is distributed across different modules. This mapping of concepts in the mind to elements in the code base of a system can be considered as central element to the comprehension of programs [5]. Thereby, not all programming paradigms enable a perfect separation [1].

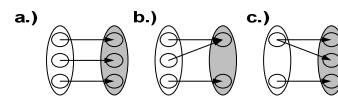


Figure 1: Mapping of the problem space to a solution space

Our hypothesis is that the concern space in the mind of a developer is not clearly separated as it is shown in Figure 1. We see support of this hypothesis, in the “strawberry are nuts” example, describing the imperfect separation and association of categories. Therefore the mapping of program elements to category concepts in the cognition is clearly problematic, since these concepts are imperfectly separated themselves. We believe that the term concern is in cognitive science, the term category. Software is created by humans, therefore the way how humans structure information, internally needs to be considered. Out of this reason we see the necessity, to match concerns and categories and discuss this issue. We think the research about categorization needs to be investigated and contains ramifications for SOC. We believe a new program comprehension model based on categorization can be created. Also, if there are a kind of family resemblances in source code needs to be considered; share elements, belonging to a concern, a family resemblance? Maybe such resemblances can be revealed by algorithms and be used to detect concerns automatically. Finally, the question, if Bananas are “more” berries than strawberries and its impact for SOC needs to be discussed.

4. REFERENCES

- [1] W. L. Hirsch, C.V.Lopes. Separation of Concerns. College of Computer Science, Northeastern University Boston, 1995
- [2] F. G. Ashby, E.M.Waldron, et.al.. Suboptimality in human categorization and identification. Journal of Experimental Psychology, 2001
- [3] D. L. Medin, E. J. Heit. Categorization. In D. Rumelhart and B. Martin Handbook of cognition and perception. Academic Press. 1999
- [4] E. W. Dijkstra. Selected Writings on Computing: A Personal Perspective. On the role of scientific thought. Springer. 1982
- [5] B.Cleary, C. Exton, Assisting Concept Location in Software Comprehension, 19th Annual Workshop of the Psychology of Programming Interest Group, Finland. 2007
- [6] S. M. Sutton Jr., P. Tarr. Aspect-oriented design needs concern modeling AOSD 2002.
- [7] E. Rosch, C.B.Mervis, Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 1975
- [8] S.P. Nguyen, G. L. Murphy. An apple is more than a fruit: Cross-classification in children’s concepts. Child Development, 74, 2003