Proceedings of the

QR'18: 31st International Workshop on Qualitative Reasoning

Co-located at Int'l Joint Conference on Artificial Intelligence (IJCAI'18)
Stockholm, Sweden, 13 July 2018

Edited by

Zoe Falomir
Bremen Spatial Cognition Centre (BSCC)
University of Bremen, Germany

and

George M. Coghill and Wei Pang
University of Aberdeen, United Kingdom

Workshop Chairs
Zoe Falomir, University of Bremen, Germany
George M. Coghill, University of Aberdeen, UK
Wei Pang, University of Aberdeen, UK

Scientific Committee
Núria Agell, ESADE - Ramon Llull University, Spain
José V. Álvarez, Universidad de Valladolid, Spain
Elena Andonova, New Bulgarian University, Bulgaria
Gautam Biswas, Vanderbilt University, USA
Mehul Bhatt, Örebro University, Sweden; and University of Bremen, Germany
Stefano Borgo, ISTC CNR, Italy
Ivan Bratko, University of Ljubljana, Slovenia
Elizabeth Bradley, University of Colorado, USA
Bert Bredeweg, University of Amsterdam, The Netherlands
Maria Chang, IBM, USA
Tony Cohn, University of Leeds, UK
Juan J. Flores, Universidad Michoacana, Mexico
Ken Forbus, Northwestern University, USA
Christian Freksa, University of Bremen, Germany
Scott Friedman, Smart Information Flow Technologies, USA
Joanna Golińska-Pilarek, University of Warsaw, Poland
Luis Gonzalez-Abril, Universidad de Sevilla, Spain
Florian Grigoleit, Technical University München, Germany
Hans Guesguen, Massey University, New Zealand
Tomoya Horiguchi, Kobe University, Japan
Liliana Ironi, IMATI - CNR, Italy
Kamal Kansou, Centre de recherche Angers-Nantes, France
Alexandra Kirsch, University of Tübingen, Germany
Johan de Kleer, Palo Alto Research Center, USA
Matthew Klenk, Palo Alto Research Center, USA
Kate Lockwood, University of St. Thomas, USA
Andrew Lovett, Northwestern University, USA
Emilio Muñoz-Velasco, University of Malaga, Spain
Lledó Museros, University Jaume I, Spain
Chris Price, Aberystwyth University, Wales, UK
Mónica Sánchez, Universitat Politècnica de Catalunya, Spain
Paulo Salles, University of Brasília, Brazil
Paulo Santos, Centro Universitario FEI, Brasil
Qiang Shen, Aberystwyth University, Wales, UK
Carl Schultz, Aarhus University, Denmark
Steven Schockaert Cardiff University, UK
Peter Struss, Technical University München, Germany
Kazuko Takahashi, Kwansei Gakuin University, Japan
Stefania Tentoni, IMATI - CNR, Pavia, Italy
Louise Travé-Massuyès, LAAS-CNRS, France
Przemysław Wałęga, University of Oxford, University of Warsaw.
Jan O. Wallgrün, Penn State University, USA
Diedrich Wolter, University of Bamberg, Germany
Jure Zabkar, University of Ljubljana, Slovenia

Preface

This book contains the accepted papers at 31^{st} *International Workshop on Qualitative Reasoning (QR'18)*[1]. QR'18 was co-located at the Int'l Joint Conference on Artificial Intelligence (IJCAI'18) and held in Stockholm, on 13^{th} July 2018.

These QR'18 proceedings contains 11 accepted papers that were presented at the workshop. Each submitted paper was reviewed by two/three program committee members. Moreover, there were 2 keynotes and 1 invited talk whose abstracts are also included in these proceedings.

The Qualitative Reasoning (QR) community develops qualitative representations to understand the world from incomplete, imprecise, or uncertain data. Our qualitative models span natural systems (e.g., physics, biology, ecology, geology), social systems (e.g., economics, cultural decision-making), cognitive systems (e.g., conceptual learning, spatial reasoning, intelligent tutors, robotics), and more.

The QR community includes researchers in Artificial Intelligence, Engineering, Cognitive Science, Applied Mathematics, and Natural Sciences, commonly seeking to understand, develop, and exploit the ability to reason qualitatively. This broadly includes:

- Developing new formalisms and algorithms for qualitative reasoning.
- Building and evaluating predictive, prescriptive, diagnostic, or explanatory qualitative models in novel domains.
- Characterizing how humans learn and reason qualitatively about the (physical) world with incomplete knowledge.
- Developing novel, formal representations to describe central aspects of our world: time, space, change, uncertainty, causality, and continuity.

The International Workshop on QR provides a forum for researchers from multiple perspectives to share research progress toward these goals. Topics of interest include:

- Qualitative modeling in physical, biological and social sciences, and in engineering.
- Representations and techniques for QR.
- Methods that integrate QR with other forms of knowledge representation, including quantitative methods, machine learning and other formalisms.
- Using QR for diagnosis, design, and monitoring of physical systems.
- Applications of QR, including education, science, and engineering.
- Cognitive models of QR, including the use of existing QR formalisms for cognitive modeling and results from other areas of cognitive science for qualitative reasoning.
- Using QR in understanding language, decision-making, sketches, images, and other kinds of signals and data sources.
- Formalization, axiomatization, and mathematical foundations of QR.

[1] QR'18: `http://homepages.abdn.ac.uk/pang.wei/pages/QR2018/index.html`

Acknowledgements

The funding from the University of Bremen and the University of Aberdeen is gratefully acknowledged. The support from the project *Cognitive Qualitative Descriptions and Applications*[2] (CogQDA) funded by the University of Bremen is also acknowledged.

We would like also to thank the members of the Scientific Committee for their valuable work during the reviewing process and the additional reviews. We also thank Easychair, which was used to manage paper submissions and reviewing the proceedings.

Zoe Falomir
George M. Coghill
Wei Pang
QR'18 Co-chairs
July 2018

[2] CogQDA: `https://sites.google.com/site/cogqda/`

Table of Contents: QR'18

Humanly AI: Creating smart people with Qualitative Reasoning

Bert Bredeweg
Faculty of Science Informatics Institute, University of Amsterdam
B.Bredeweg@uva.nl

Abstract

Modern stochastic-based AI is without doubt responsible for many of the recent successes involving machine learning and pattern recognition. This swing back to Goofy control theory and cybernetics is good for industry and owners of big data. But what does it entail for people? End of privacy, autonomous weapons, fake news?

We prefer to take a very different approach, and are interested in developing AI that supports people, not destroys them. Hence, we develop Humanly AI, which is articulate, reflective and communicable. Articulate refers to being explicit, either in knowledge (knowing things) or in formats (knowing how to represent things). Reflective refers to being able to assess and process what is articulated (e.g. compare, contrast, order, etc.). Communicable refers to being able to interactively share and co-construct, typically with the goal to obtain more and better knowledge, possible for the AI but surely also for the human in the loop.

This presentation will illustrate the added value of Qualitative Reasoning for developing Humanly AI and how this can be deployed for learners in secondary education (e.g. Schlatter et al., EC-TEL 2017) and for scientists doing cutting edge research (e.g. Kansou et al., Scientific Reports 2017). We argue that this is more fun, more valuable, and much less dangerous then Goofy control theory and cybernetics.

References

Schlatter, E., Bredeweg, B., Drie, J. Van, and Jong, P. de. (2017). Can Learning by Qualitative Modelling Be Deployed as an Effective Method for Learning Subject-Specific Content? In: É. Lavoué et al. (Eds.): EC-TEL 2017, LNCS 10474, pp. 479-485.

Kansou, K., Rémond, C., Paës, G., Bonnin, E., Tayeb, J. and Bredeweg, B. (2017). Testing scientific models using Qualitative Reasoning: Application to cellulose hydrolysis. Scientific Reports 7, 14122.

About the Keynote speaker

Dr. Bert Bredeweg is an associate professor at the Informatics Institute, within the Faculty of Science of the University of Amsterdam. His research concerns computational intelligence and includes: knowledge capture, qualitative reasoning, learning by modelling, cognitive diagnose, and human-computer interaction. Bredeweg has published over 139 international refereed academic publications, including 38 journal and 77 proceedings papers, 12 book chapters and 12 edited collections. These publications have found their way into leading journals, and competitive conferences (he also produced over 137 other academic publications). Bredeweg has supervised over 70 MSc and PhD students. He is an established and active member of the Ecological informatics, Qualitative Reasoning, and Educational technology communities. He is a regular senior reviewer and board member for the associated conferences and journals.

Spatial Reasoning: Theory, Methods and Applications

Mehul Bhatt
Örebro University, Sweden; and University of Bremen, Germany
mehul.bhatt@oru.se; bhatt@uni-bremen.de

Abstract

This talk presents spatial representation and reasoning from the viewpoints of the research areas of knowledge representation and reasoning, commonsense reasoning, and spatial cognition and computation. The central focus of the talk is on "declarative spatial reasoning” denoting: (1). the ability to `declaratively' specify and solve real-world problems involving mixed qualitative-quantitative visuo-spatial representation and reasoning; and (2). systematic formalisation and (corresponding) general declarative programming methods supporting reasoning capabilities involving, e.g., query answering, inductive (relational) learning, non-monotonic abductive inference pertaining to primitives of space and motion. Towards (1-2), I will particularly emphasise the development of domain-independent methods for commonsense reasoning about space, action, and change (whilst presenting examples with semantically-driven methods rooted in constraint logic programming, inductive logic programming, and answer set programming).

Our research in spatial reasoning has been driven by specialist and everyday commonsense reasoning instances identifiable in a range of cognitive technologies and spatial assistance systems where spatio-linguistic conceptualisation & background knowledge focussed visuo-spatial cognition and computation are central. In this backdrop, I will conclude by particularly highlighting the significance of knowledge-centred technologies focussing on commonsense semantics and qualitative reasoning & representation learning for mixed-methods research at the interface of AI, Psychology, and (Media & Architecture) Design.

More info: CoDesign Lab EU. www.codesign-lab.org and www.mehulbhatt.org

References

M. Bhatt. Reasoning about Space, Actions and Change: A Paradigm for Applications of Spatial Reasoning. In Qualitative Spatial Representation and Reasoning: Trends and Future Directions. IGI Global, USA, 2012. ISBN ISBN13: 9781616928681. (2012)

M. Bhatt and S. W. Loke. Modelling Dynamic Spatial Systems in the Situation Calculus. Spatial Cognition & Computation, 8(1-2): 86–130, 2008. doi: 10.1080/13875860801926884. URL http://dx.doi.org/10.1080/13875860801926884. (2008)

M. Bhatt, J. H. Lee, and C. P. L. Schultz. CLP(QS): A declarative spatial reasoning framework. In Spatial Information Theory - 10th International Conference, COSIT 2011, Belfast, ME, USA, September 12-16, 2011. Proceedings, pages 210–230, 2011. doi: 10.1007/978-3-642-23196-4_12. URL http://dx.doi.org/10.1007/978-3-642-23196-4_12. (2011)

M. Bhatt, C. Schultz, and C. Freksa. The ‘Space’ in Spatial Assistance Systems: Conception, Formalisation and Computation. In: Representing space in cognition: Interrelations of behavior, language, and formal models. Series: Explorations in Language and Space. 978-0-19-967991-1, Oxford University Press, 2013. (2013)

J. Suchan, M. Bhatt, and C. P. L. Schultz. Deeply semantic inductive spatio-temporal learning. In J. Cussens and A. Russo, editors, Proceedings of the 26th International Conference on Inductive Logic Programming (Short papers), London, UK, 2016., volume 1865, pages 73–80. CEUR-WS.org, 2016. URL http://ceur-ws.org/Vol-1865/paper-12.pdf. (2016)

P. A. Walega, M. Bhatt, and C. P. L. Schultz. ASPMTQS(QS): Non-monotonic spatial reasoning with answer set programming modulo theories. In Logic Programming and Nonmonotonic Reasoning - 13th International Conference, LPNMR 2015, Lexington, KY, USA, September 27-30, 2015. Proceedings, pages 488–501, 2015. doi: 10.1007/978-3-319-23264-5_41. URL http://dx.doi.org/10.1007/978-3-319-23264-5_41. (2015)

J. Suchan and M. Bhatt. Semantic question-answering with video and eye-tracking data: AI foundations for human visual perception driven cognitive film studies. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016, pages 2633–2639, 2016. URL http://www.ijcai.org/Abstract/16/374. (2016)

J.Suchan, M. Bhatt, P. Walega,and C. Schultz. Visual Explanation by High-Level Abduction: On Answer-Set Programming driven Reasoning about Moving Objects. In AAAI 2018 : Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, 2018. (2018)

About the Keynote speaker

Mehul Bhatt is Professor within the School of Science and Technology at Örebro University (Sweden) and Guest Professor at the Department of Computer Science at the University of Bremen (Germany). Mehul's research encompasses the areas of artificial intelligence, spatial cognition and computation, visual perception, and human-computer interaction. His research addresses applications in the fields of autonomous systems, architecture design, visuo-auditory media design, and geospatial systems.

Mehul Bhatt directs the research and consulting group DesignSpace (www.design-space.org), and has most recently launched CoDesign Lab EU (www.codesign-lab.org), an initiative aimed at addressing the confluence of Cognition, AI, Interaction, and Design Science. Previously, he led the Human-Centred Cognitive Assistance Lab (http://hcc.uni-bremen.de) at the University of Bremen. Mehul Bhatt can be reached via: www.mehulbhatt.org

Integrating Qualitative Reasoning and Learning: Prospects and Problems

Diedrich Wolter
University of Bamberg, Germany
diedrich.wolter@uni-bamberg.de

Abstract

A common motivation of qualitative reasoning is to prepare computers to solve everyday problems that require a basic understanding of environmental sciences, for example physical problems of stability when stacking up dishes. With 'prepare to solve' I refer to the symbolic nature of qualitative reasoning which is involved with representations for modeling the problem domain and methods for reasoning about actions and their effects, rather than, say, stacking up dishes. Qualitative reasoning aims to match human problem solving skills and has in some tasks already exceeded those. However, a particular feature still sets apart human from computer problem solving: the superior ability of humans to learn, to adapt to new circumstances, and, last but not least, to reflect about what has been learnt. In this talk I discuss approaches aiming to combine learning with qualitative reasoning. The presentation is aimed at identifying opportunities of integrating learning and reasoning, as well as discussing open research questions for the QR community.

About the Invited Speaker

Diedrich has graduated 2006 from the university of Bremen with a dissertation on combining computer vision techniques and spatial reasoning to tackle the self-localisation problem of mobile robots. As a principal investigator in the spatial cognition research center he and his team worked on qualitative spatial reasoning, developing new representations, and applying them to mobile robots. He joined Bamberg university in 2013 as professor for smart environments where he still works on qualitative spatial reasoning and investigates how these techniques can be applied in various application domains.

Qualitative Shape Representation and Reasoning Based on Concavity and Tangent Point

Momo Tosue and **Sosuke Moriguchi** and **Kazuko Takahashi**
Kwansei Gakuin University
eyi98098@kwansei.ac.jp chiguri@acm.org ktaka@kwansei.ac.jp

Abstract

We propose a method for qualitative analysis of changes in the shape of an object, focusing on the generation of concavities, tangent points, and division. We aim at qualitative treatment on changes in the shape of a cell sheet arising during organogenesis. We develop a descriptive language that qualitatively represents the shape of an object at a low level, and then a method extracting features of the object from the expression in this language. To afford a higher-level representation, we classify shapes using only the concavity and tangent point status, and the number of components. We develop transition rules governing these qualitative shape representations and show the state transitions. This enables both qualitative simulation and backward reasoning when an unexpected state arises.

1 Introduction

Recently, life science has become an important research field. It is very interesting to analyze or formalize processes in the developmental biology from a viewpoint of computer science. At the same time, such investigations will contribute the advancement of life science.

Organogenesis commences with a sphere termed an alveolus, the surface of which is covered with a sheet of cells. This sheet changes in shape via simple transformations such as folding or splitting, gradually becoming an organ. Thus, eyes, ears, and neural tubes are formed via diverse shaping of cell sheets. Although these developmental changes are continuous, it is reasonable to create a qualitative model to analyze the principal changes and their causes. In a qualitative simulation, we consider only the states at which major events occur and the transitions between them [de Kleer, 1993; Kuipers, 1993; Forbus, 2010]. However, qualitative simulation has not yet been applied to deal with changes in shape such as folding or splitting.

Let us consider an example. Figure 1 shows the organogenesis of an eye based on [Wolpert and Tickle, 2011]. The part highlighted in bold, termed the crystalline lens plate, will transform into the eye. In the states of (A) and (B), the plate is not bent, but then becomes bent, generating a concavity, when developing into state (C), and a new object termed a lens cell separates from the plate in state (D) after the entrance of the concavity is closed.

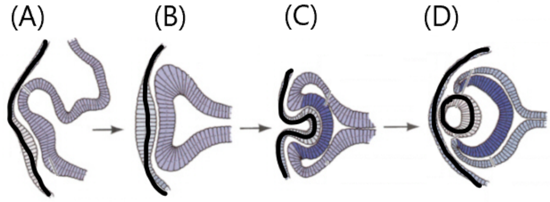


Figure 1: Organogenesis of an eye.

This can be modeled as a qualitative shape change of the cell sheet, as shown in Figure 2.

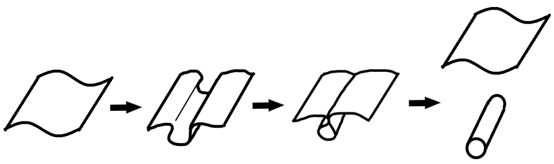


Figure 2: A model for organogenesis of an eye.

In this paper, we take an approach of Qualitative Spatial Reasoning (QSR) toward the representation and reasoning on changes in shape during organogenesis. QSR is a method used for representation and analysis without using precise numerical data. It focuses on certain aspects of an object, and reasons by reference to those aspects. It has an advantage mainly on the following points: it does not require extensive computational resources and it fits human cognition.

Over the past 30 years, many works on QSR have been published [Stock, 1997; Cohn and Hazarika, 2001; Cohn and Renz, 2007; Ligozat, 2011; Chen et al., 2013]. of which several focused on the shape representation of objects projected onto a two-dimensional plane [Leyton, 1988; Cohn, 1995; Schlieder, 1996; Galton and Meathrel, 1999; Kulik and Egenhofer, 2003; Gottfried, 2003; Gottfried, 2004; Museros and Escrig, 2004]. These objects generally featured closed boundaries and had neither an end point nor a tangent point. On the other hand, a cell sheet may connect to itself, creating a tangent point, and the sheet may be cut at that point. Such a change in shape cannot be represented using existing methods.

In this paper, we first develop a language representing the shape of an object. Our initial low-level representation employs a directed line and a point as primitive terms, providing a qualitative picture of the outline of an object. Then, we extract the characteristics of the shape from this representation. These characteristics might include the existence of concavities and/or tangent points. We thus generate a high-level abstract representation; we use predicates to this end. Next, we develop state transition rules applicable at the higher level. Using these rules, we show that a qualitative change frequently found during organogenesis can be modeled.

Moreover, if we impose certain transition rules on the symbolic representation, new characteristics may be extracted and a new, high-level state transition rule may be generated. As a result, we may find an unexpected state and determine the reason why it appeared.

This paper is organized as follows. After describing some basic concepts of graph theory and elementary geometry in Section 2, we describe shapes and the changes to be modeled in Section 3. In Section 4, we develop a language representing the shape of an object at a low level. Then, we extract the features of the shape from this expression. In Section 5, we construct a state transition system employing these features. In Section 6, we compare our work with related works and in Section 7, we show conclusions and mention our future works.

2 Preliminaries

We here summarize several basic concepts of graph theory [Harary, 1969] and elementary geometry.

A graph is defined as a pair of a set of vertices and a set of edges. For a graph, a sequence of edges $E_1, \ldots, E_n$ with E_i connecting vertices V_{i-1} and V_i ($1 \leq i \leq n$) is said to be *a walk*, and if $V_0 = V_n$, the walk is said to be *closed*. For a walk, if $E_i \neq E_j$ for each pair of (i, j) ($1 \leq i < j \leq n$), the walk is said to be *a trail*; if $V_i \neq V_j$ for each pair of (i, j) ($0 \leq i < j \leq n$), the walk is said to be *a path*. A closed trail is said to be *a circuit*. For a closed trail, if $V_i \neq V_j$ for each pair of (i, j) ($0 \leq i < j \leq n-1$), then it is said to be *a cycle*. A walk that visits every edge exactly once is said to be *an Eulerian trail*. For a vertex, the number of outgoing or incoming edges is said to be *the degree* of the vertex. If zero or two vertices in a graph are of odd degree, then the graph has an Eulerian trail. An Eulerian trail that starts and ends on the same vertex is said to be *an Eulerian circuit*. A graph containing an Eulerian circuit is said to be *an Eulerian graph* (Figure 3(a)). A graph containing an Eulerian trail but not an Eulerian circuit is said to be *a semi-Eulerian graph* (Figure 3(b)).

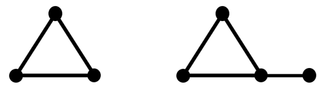


Figure 3: (a) An Eulerian graph and (b) a semi-Eulerian graph.

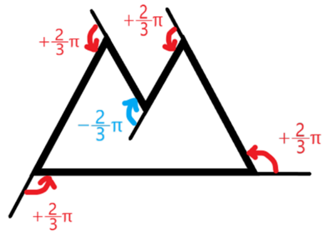


Figure 4: Exterior angle.

The outline of a polygon can be considered to be a graph. Here, we determine the direction of each edge by starting from an arbitrary vertex and tracing the boundary anticlockwise, viewing from the left. Let E_1 be an edge and E_2 be the adjacent edge of a polygon. Then, the *exterior angle* of a vertex is defined as the angle made by the extension of E_1 and the edge E_2. When the vertex is convex, the exterior angle is positive, while when the vertex is concave, the angle is negative. The sum of the exterior angles of a polygon is 2π (Figure 4).

We recall the following propositions from elementary geometry.

Proposition 1 *Let $E_1, \ldots, E_n$ be a sequence of edges, and θ_i ($1 \leq i \leq n$) be the exterior angle between E_{i-1} and E_i. Then, an Eulerian graph can be drawn iff $\Sigma_{1 \leq i \leq n} \theta_i = 2\pi$.*

Proposition 2 *An Eulerian trail can be drawn without an intersection.*

It follows immediately that an Eulerian graph and a semi-Eulerian graph can be drawn using one stroke.

3 Shapes and changes to be modeled

We consider the shape of the cross-section of a cell sheet projected onto the two-dimensional plane.

Generally, organogenesis features three important transformations of a cell sheet: concavity generation, tangent point generation, and division. Therefore, we classify shapes depending on these features. Shapes that exhibit the same characteristics as the concavity and the tangent point are regarded as equivalent. Objects with and without end points can be handled. The orientation of a figure, its size, and the number of concavities, are ignored. We show examples of shapes which may appear by transforming a cell sheet in Figure 5.

In this figure, the shapes within the dotted rectangles are regarded as the same, whereas shapes in different rectangles are considered different.

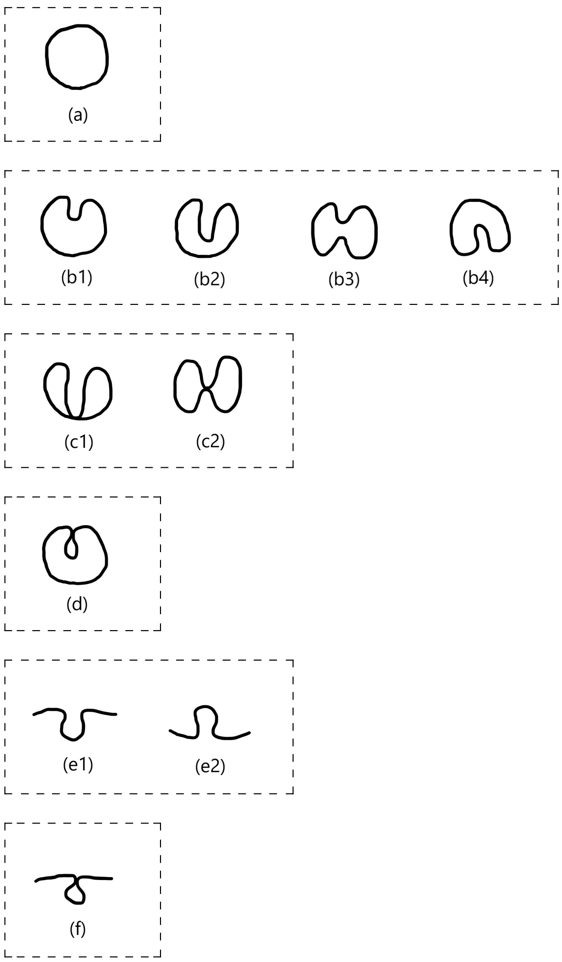


Figure 5: Examples of shapes appearing in the organogenesis process.

In fact, the ends of a cell sheet are connected to an organ but, as we investigate local changes in shape, we assume that the sheet is of finite size. A cell sheet never crosses itself, because it is a "sheet." It means that the figure can be drawn using one stroke without an intersection.

Figure 6 shows one of changes that frequently appears in the organogenesis process, to which we are going to give a qualitative representation.

4 Description Language

4.1 Language

We develop a language $\mathcal{L}$ representing the shape of the outline of an object.

$exp ::= (seg+) \mid [seg+]$
$seg ::= dline \mid point$

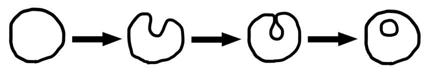


Figure 6: An example of a shape change appearing in the organogenesis process.

$dline ::= r \mid r^+ \mid r^- \mid l \mid l^+ \mid l^-$
$point ::= a \mid b \mid \ldots$ where $a, b, \ldots$ are constants.

To eliminate redundancy, we impose the constraint that a segment never appears immediately after the same segment.

The language $\mathcal{L}$ is a set of expressions exp, satisfying this constraint. exp is a sequence of segments surrounded by parentheses or brackets; if an exp is surrounded by parentheses, then exp is called *an open expression*; and if it is surrounded by brackets, then exp is called *a closed expression*. The sequence between parentheses or brackets is called *a sequence of the expression*. Segment seg is either $dline$ or $point$. There are six types of $dline$, and $point$ is a constant.

S, D, and P, denote the sets of all segments, all directed lines, and all points, respectively. Then, $S = D \cup P$ and $D \cap P = \emptyset$ hold.

The structure of a closed expression $[x_1 \ldots x_n]$ is cyclic, which means that $x_{i+n} = x_i$ holds for any i $(0 < i \leq n)$. For an open expression $(x_1 \ldots x_n)$, x_{i+n} $(0 < i \leq n)$ is undefined, denoted as $\perp$.

We introduce a function $succ$. For a sequence $x_1 \ldots x_n$, $x_{succ(i)}$ indicates the $dline$ that appears after x_i.

$$succ(i) = \begin{cases} i+1 & (\text{if } x_{i+1} \in D) \\ i+2 & (\text{if } x_{i+1} \in P) \\ \perp & (\text{if } x_{i+1} \text{ is } \perp) \end{cases}$$

For an expression e and a point p, $occur(e, p)$ indicates the number of times p occurs in e.

4.2 Semantics

An expression corresponds to a figure drawn on a two-dimensional plane. Intuitively, an expression is the trace of an outline of an object. A closed expression corresponds to a case in which the start point and the end point coincide, whereas an open expression corresponds to a case in which the points are different. $dline$ means a directed line that has $\pi/3$ steps, the lengths of which are ignored (Figure 7), and $point$ indicates an intersection.

We define the angle of rotation between two segments as a function rot from $S \times S$ to $\{n\pi/3 \mid n \in \{-2, -1, 0, 1, 2, 3\}\}$[1].

- For $d \in D$,
  $rot(r, r) = 0$. $\quad rot(r, r^+) = \pi/3$.
  $rot(r, l^+) = 2\pi/3$. $\quad rot(r, l) = \pi$.
  $rot(r, l^-) = -2\pi/3$. $\quad rot(r, r^-) = -\pi/3$.
  For the other $d, d' \in D$, $rot(d, d')$ is similarly defined.
- For $p \in P$ and $x \in S$,

[1]Here, we use a global coordinate axis to describe the outline, but a description using a relative axis is also available.

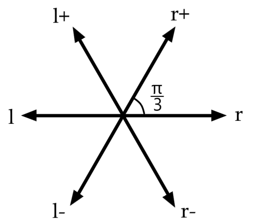


Figure 7: Directed lines.

– $rot(p, x) = 0. \quad rot(x, p) = 0.$

For a sequence $s = x_1 \ldots x_n$, if $\Sigma_{i=1}^{n} rot(x_i, x_{succ(i)}) = \pm 2\pi$, then s is said to be a *closed sequence*.

A function rev from D to D is defined as follows:

- For $d \in D, \quad rev(d) = d'$ iff $rot(d, d') = \pi$.

Intuitively, $rev(d)$ is a directed line in the direction opposite to d. Clearly, $rev(rev(d)) = d$ holds for any $d \in D$.

The language $\mathcal{L}$ can represent all shapes formed by the transformation of a cell sheet. For example, let $exp_1 = (r\ a\ l^-\ r\ l^+\ a\ r)$ and $exp_2 = [r\ l^+\ l^-\ l^+\ l^-]$ be expressions. Then, their corresponding figures are shown in Figure 8 and Figure 9, respectively.

(→ • ↙ → ↖ • →)

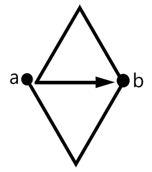
Figure 8: A figure corresponding to exp_1.

[→ ↖ ↙ ↖ ↙]

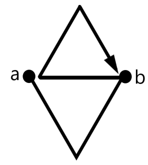
Figure 9: A figure corresponding to exp_2.

It is interesting to note that different expressions correspond to the same figure. Figure 10 is a figure equivalent to both expressions: $exp_3 = (a\ r^-\ r^+\ b\ l^+\ l^-\ a\ r\ b)$ and $exp_4 = (a\ r^-\ r^+\ b\ l\ a\ r^+\ r^-\ b)$.

Similarly, Figure 11 corresponds to both expressions: $exp_5 = (r^-\ a\ l\ r^-\ r^+\ l\ a\ r^-\ l\ r^+\ a\ r^+)$ and $exp_6 = (r^-\ a\ l\ r^-\ r^+\ l\ a\ l^-\ r\ l^+\ a\ r^+)$. However, if we draw this figure according to exp_6, we have to cross lines, which cannot be achieved by transformation of a sheet. Therefore, we will later impose conditions on an expression to eliminate such a case.

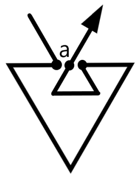


(a) Drawn by reference to exp_3.

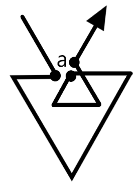


(b) Drawn by reference to exp_4.

Figure 10: A figure corresponding to exp_3 and exp_4.



(a) Drawn by reference to exp_5.



(b) Drawn by reference to exp_6.

Figure 11: A figure corresponding to exp_5 and exp_6.

4.3 Consistent expression

We discuss the conditions an expression should satisfy so that a corresponding figure that can be obtained by transformation of a sheet exists.

First, we introduce the set $Split$. For an expression e and a point p, $Split(e, p)$ is defined to be a set of closed sequences, each of which is surrounded by the same point, respectively.

$$Split(e, p) = \{x_i \ldots x_j \mid 1 < i < j \wedge j - i < n \wedge x_{i-1} = x_{j+1} = p\}$$

For example, $Split(exp_5, a) = \{l\ r^-\ r^+\ l,\ r^-\ l\ r^+,\ l\ r^-\ r^+\ l\ a\ r^-\ l\ r^+\}$, and $Split(exp_6, a) = \{l\ r^-\ r^+\ l,\ l^-\ r\ l^+,\ l\ r^-\ r^+\ l\ a\ l^-\ r\ l^+\}$.

The following conditions ensure that segments are never superposed on the plane.

Definition 1 *Let e be an expression of which sequence is $x_1 \ldots x_n$. If e satisfies all of the following conditions, then it is called* a consistent expression.

1. $\forall i (\forall p, p' \in P(x_i = p \Rightarrow x_{i+1} \neq p'))$.
2. $\forall i (\forall d \in D(x_i = d \Rightarrow x_{i+1} \neq rev(x_i)))$.
3. *For $p \in P$, let $I(e, p) = \{i - 1 \mid 0 < i \leq n \wedge x_i = p\}$ and $O(e, p) = \{i + 1 \mid 0 \leq i < n \wedge x_i = p\}$.*
   - $\forall p \in P(\forall in \in I(e, p), out \in O(e, p), x_{in} \neq x_{rev(out)})$.
   - $\forall p \in P(\forall in, in' \in I(e, p), \forall p \in P(\forall in, in' \in I(e, p), in \neq in' \Rightarrow x_{in} \neq x_{in'})$.
   - $\forall p \in P(\forall out, out' \in O(e, p), \forall p \in P(\forall out, out' \in O(e, p), out \neq out' \Rightarrow x_{out} \neq x_{out'})$.
4. $\forall p \in P(occur(e, p) \neq 1)$.
5. *If $e = (x_1 \ldots x_n)$, then $\forall p \in P(x_1 \neq p \vee x_n \neq p)$.*
6. *If $e = [x_1 \ldots x_n]$, then $x_1 \ldots x_n$ is a closed sequence.*

7. $\forall p \in P(\forall e' = x_1 \ldots x_n \in Split(e,p),$
   $x_1 \ldots x_n$ *is a closed sequence).*

Condition 3 means that directed lines to/from a point are never crossed.

Condition 7 is imposed so that the corresponding figure can be drawn without crossing lines. For example, as each element of $Split(exp_5, a)$ is a closed sequence, exp_5 satisfies condition 7; on the other hand, as the first two elements of $Split(exp_6, a)$ are closed sequences, whereas the third is not, exp_6 does not satisfy condition 7.

For an expression in $\mathcal{L}$, we can easily obtain the corresponding graphical expression by regarding each directed line as an edge and the start and end points as vertices, respectively. Thus, from Proposition 1 the following proposition holds:

Proposition 3 *For a consistent expression e, there exists an Eulerian trail.*

It follows from Proposition 2 that a consistent expression e can be drawn without an intersection.

5 Higher-Level Analysis

5.1 Extraction of a feature

We generate an abstract representation from an expression in $\mathcal{L}$. In this higher-level analysis, we represent only the characteristics of shapes based on concavities and tangent points, together with the number of included cycles, whereas $\mathcal{L}$ represents a configuration of segments. A figure should be evaluated with respect to these three features. We ignore the relative extents of concavities and curvatures, and classify shapes using only the existence or not of a concavity and/or a tangent point, and the number of cycles.

Let e be a consistent expression. $D(e), T(e)$ and $N(e,k)$ indicate that e has a concavity, e has a tangent point to itself, and e has k cycles, respectively. We can omit e if the expression is trivial.

- $D(e)$ holds if $\exists i; (rot(x_i, x_{i+1}) \times rot(x_{i+1}, x_{i+2}) < 0)$
- $T(e)$ holds if $\exists p \in P; occur(e,p) \neq 0$
- $N(e,k)$ holds if the number of cycles of e is k when we regard the expression as a graph. Formally, k is determined as follows:

$$k = \begin{cases} \Sigma_{p \in P} max(occur(e,p) - 1, 0) \\ \qquad\qquad (\text{if } e = (x_1 \ldots x_n)) \\ \Sigma_{p \in P} max(occur(e,p) - 1, 0) + 1 \\ \qquad\qquad (\text{if } e = [x_1 \ldots x_n]) \end{cases}$$

For a consistent expression e, the characteristics can be represented using these predicates. There are four possible combinations of the truth values of D and T, but $D \wedge T$ can be omitted from consideration, as the entrance to a concave part is closed and a concavity disappears when a tangent point appears.

5.2 State transition

Next, we construct state transition rules using the predicates D, T and N.

The following properties hold: when a tangent point is generated, the number of cycles increases; and, when division occurs, the number of cycles are divided between the two resulting objects.

Thus, we obtain the following three state transition rules. The symbol '$\rightarrow$' indicates a direct transition that is *a conceptual neighbor* in the case of a single direction [Freksa, 1992].

For a consistent expression e:

- (R1) Generation of a concavity
  $\neg D(e) \wedge \neg T(e) \wedge N(e,k) \rightarrow D(e) \wedge \neg T(e) \wedge N(e,k)$
- (R2) Generation of a tangent point
  $D(e) \wedge \neg T(e) \wedge N(e,k) \rightarrow \neg D(e) \wedge T(e) \wedge N(e,k+1)$
- (R3) Division
  $\neg D(e) \wedge T(e) \wedge N(e,k) \rightarrow (\neg D(e_1) \wedge \neg T(e_1) \wedge N(e_1,k_1)) \wedge (\neg D(e_2) \wedge \neg T(e_2) \wedge N(e_2,k_2))$
  where $k = k_1 + k_2, k_1, k_2 \geq 0$

It is sufficient to consider the case for $n = 0, 1, 2$ in terms of the state transition rules. There are three constraints on state transitions: (i) when a concavity is generated, a tangent point is not generated at the same time, (ii) a tangent point is never generated without generation of a concavity, and (iii) an object without a tangent point never divides. Therefore, we have six classes of possible shapes, as shown in Table 1.

Table 1: Possible shapes.

feature \ N(k)	k=0	k=1	k=2
¬D∧¬T			
D∧¬T			
¬D∧T			

The class $\neg D \wedge T \wedge N(2)$ has two shapes. We use one of them (for example, the former) as a representative of the class. The expressions in $\mathcal{L}$ for representatives of each class are shown in Figure 12. It is easy to check that individual features can be extracted from these expressions, respectively.

Applying the above transformation rules, we obtain the state transition graphs shown in Figure 13. Note that when an object divides, we represent only the shape of each object formed, ignoring their relative positions. Such changes may be noted during most organogeneses, such as those of the lens of the eye, the semicircular canal that is transformed to the inner ear tube, and the neural tube. Thus, this state transition graph affords a qualitative model of organogenesis.

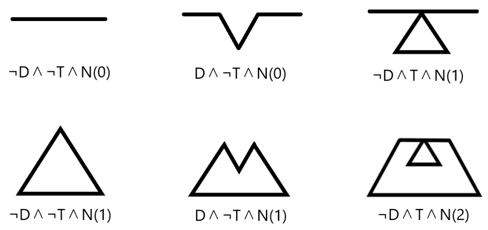


Figure 12: The expressions in $\mathcal{L}$ for each state.

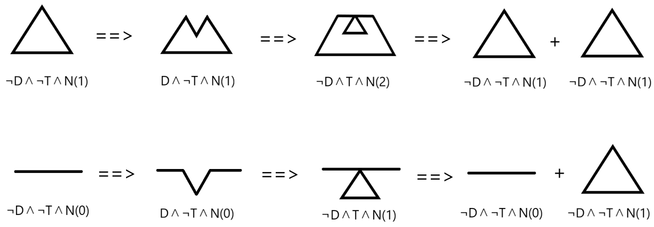


Figure 13: A qualitative model for an organogenesis process.

5.3 Granularity refinement

The class $\neg D \wedge T \wedge N(2)$ has two shapes that differ both topologically and cognitively. It is natural that these should be discriminated. Both shapes can be outcomes of a change from the class $D \wedge \neg T \wedge N(1)$. One is obtained by closing the entrance of the concavity to an external tangent, whereas the other is obtained by deepening the concavity to internalize the tangent (Figure 14).

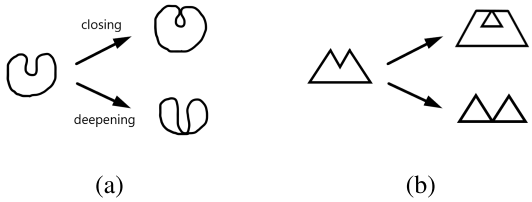


Figure 14: (a) Two different shapes of $\neg D \wedge T \wedge N(2)$ and (b) their expressions in $\mathcal{L}$.

These two figures are associated with different expressions in $\mathcal{L}$, whereas both are in the same class $\neg D \wedge T$ at the higher level. This means that the abstraction is too coarse. Therefore, we should assign intermediate granularity to the representation by introducing predicates such as external and internal tangents.

Moreover, if we consider the change of shape indicated by the expression $\mathcal{L}$, new characteristics may be extracted and a new state transition rule generated. For example, we can consider a new state transition rule for the shape change shown in Figure 15. This corresponds to the change from a closed expression $[x_1 \ldots x_n]$ to an open expression $(x_1 \ldots x_n)$ in $\mathcal{L}$. We expect that neural tube obstructions arise in this way, and that other abnormal states may develop in a similar manner.

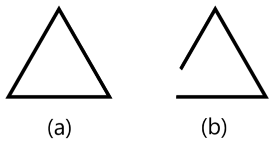


Figure 15: Possible state transition rule: (R4) Cutting.

We can describe this transition rule as follows:

- (R4) Cutting
  $\neg D(e) \wedge \neg T(e) \wedge N(e,k) \rightarrow \neg D(e) \wedge \neg T(e) \wedge N(e,k-1)$

This seems to be extended to the following rules:

- (R4') Cutting_2
  $D(e) \wedge \neg T(e) \wedge N(e,k) \rightarrow D(e) \wedge \neg T(e) \wedge N(e,k-1)$
- (R4") Cutting_3
  $\neg D(e) \wedge T(e) \wedge N(e,k) \rightarrow \neg D(e) \wedge T(e) \wedge N(e,k-1)$

However, (R4") cannot be accepted, as the figure yielded by the rule is not permitted to be a shape that a cell sheet assumes (Figure 16). The figure must be viewed as an inconsistent expression.

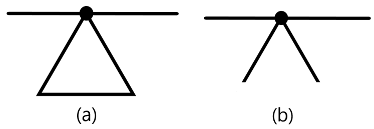


Figure 16: Unaccepted state transition rule: (R4")

.

This example indicates that we need to define a state transition in $\mathcal{L}$ or to introduce an intermediate level of abstraction.

6 Related Works

Many QSR approaches have been developed, but papers on shapes are few in number compared to those on mereological relationships or directions. The main reason is that it is difficult to identify the aspects of shapes that should be formalized.

It is natural to represent the shape of an object by tracing its boundary on a two-dimensional plane. This represents a shape as a sequence of segments, sometimes combined with the relationships between subsequent segments. Leyton developed a grammar to describe the shape of a smooth outline, based on the qualitative curvature [Leyton, 1988]. Galton and Meathrel created another shape grammar representing an outline in a similar way [Galton and Meathrel, 1999]. Unlike Leyton, the latter authors assumed that an outline consisted of a finite number of line segments.

Museros and Escrig also used line segments, but their representation additionally required a qualitative shape, an angle, or a size for each segment [Museros and Escrig, 2004]. Schlieder represented the shape of an outline via positional ordering of points on the boundary [Schlieder, 1996]. Kulik and Egenhofer developed a language to represent the characteristics of a landscape projected onto a two-dimensional plane [Kulik and Egenhofer, 2003]. They represented terrain features qualitatively using several types of primitive vectors and combinations thereof. Gottfried developed two different calculi, both of which were based on the relationships between subsequent line segments [Gottfried, 2003; Gottfried, 2004].

These languages were designed to represent only closed regions, and cannot deal with an object with an end point or a tangent point, such as that shown, for example, in Figure 8. Kulik treated a figure with end points but not a figure with a tangent point. We could extend the existing languages to represent an object with an end point or a tangent point. However, other aspects such as curvature, segment size, and concavity position that we wish to ignore are embedded in such language descriptions. Thus, we would obtain a complicated redundant representation. It seems that extensions to existing languages would not help us achieve our goal to grasp the transformation of a cell sheet, and development of a new language is a better solution.

Cohn took a different approach [Cohn, 1995], proposing a representation using relationships over regions. Convexity was considered, with a focus on the difference between the original region and its convex hull. There, subtle qualitative shape differences were represented in a hierarchical manner. The concavity and tangential point features on which we have focused can be represented by extending the his formalization. However, it would be necessary to introduce new predicates representing these features together with several axioms.

In summary, the figure used in almost all approaches had a closed boundary without a tangent point or end points. On the other hand, we allow an outline that has an end point, a tangent point, and/or a closed boundary. Another significant difference between our work and earlier papers is that we consider shape transformation including a division, while other works do not.

7 Conclusion

We have discussed a qualitative shape representation and investigated state transitions between the representations. We developed a language describing qualitative shapes in a low-level and showed a rule for extracting a higher-level representation. For a higher-level representation, we classified shapes using only the existence of a concavity, a tangent point and the number of cycles, and showed state transitions. We also discussed the granularity of expressions.

In future, we would like to define low-level state transition rules that allow us to identify hitherto unknown states. Moreover, we expect that our method can be applied not only to analysis of organogenesis but also to other applications such as exploration of change in terrain.

References

[Chen et al., 2013] Chen, J., Cohn, A., Liu, D., Wang, S., Ouyang, J., and Yu, Q. (2013). A survey of qualitative spatial representations. *The Knowledge Engineering Review*, 30(1):106–136.

[Cohn, 1995] Cohn, A. G. (1995). hierarchical representation of qualitative shape based on connection and convexity. *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science, International Conference COSIT'99 Stade, Germany, August 25-29, 1999 Proceedings*, pp. 311-326.

[Cohn and Hazarika, 2001] Cohn, A. G. and Hazarika, S. M. (2001). Qualitative spatial representation and reasoning: An overview. *Fundamental Informaticae*, 46(1-1):1–29.

[Cohn and Renz, 2007] Cohn, A. G. and Renz, J. (2007). *Handbook of Knowledge Representation*, chapter 13. Qualitative Spatial Reasoning, pp. 551–596. Elsevier.

[de Kleer, 1993] de Kleer, J. (1993). A view on qualitative physics, *Artificial Intelligence*, 59(1-2):105-114.

[Forbus, 2010] Forbus, K. D. (2010). *Qualitative modeling*. Wiley.

[Freksa, 1992] Freksa, C. (1992). Using orientation information for qualitative spatial reasoning. *Theories and methods of spatio-temporal reasoning in geographic space*, pp. 162-178.

[Galton and Meathrel, 1999] Galton, A. and Meathrel, R. (1999). Qualitative outline theory. *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pp. 1061-1066.

[Gottfried, 2003] Gottfried, B. (2003). Tripartite line tracks qualitative curvature information. *Proceedings of the Spatial Information Theory. Foundations of Geographic Information Science, International Conference, COSIT 2003*, Ittingen, Switzerland, September 24-28. Lecture Notes in Computer Science 2825, pp. 101-117. Springer.

[Gottfried, 2004] Gottfried, B. (2004). Reasoning about intervals in two dimensions. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, The Hague, Netherlands, October 10-13. IEEE, pp. 5324-5332.

[Harary, 1969] Harary, F. (1969). *Graph Theory*. Reading, Mass.: Addison-Wesley Pub. Co., 1969. Print.

[Kuipers, 1993] Kuipers, B. J. (1993). Qualitative simulation: then and now. *Artificial Intelligence*, 59(1-2):133-140.

[Kulik and Egenhofer, 2003] Kulik, L. and Egenhofer, M. J. (2003). Linearized terrain: languages for silhouette representations. *Spatial Information Theory. Foundations of Geographic Information Science, International Conference, COSIT 2003, Ittingen, Switzerland, September 24-28, 2003, Proceedings*, pp. 118-135.

[Leyton, 1988] Leyton, M. (1988). A process-grammar for shape. *Artificial Intelligence*, 34(2):213-247.

[Ligozat, 2011] Ligozat, G. (2011). *Qualitative Spatial and Temporal Reasoning*. Wiley.

[Museros and Escrig, 2004] Museros, L. and Escrig, M. (2004). A Qualitative Theory for Shape Representation and Matching for Design. *Proceedings of the Sixteenth European Conference on Artificial Intelligence, ECAI'2004, including Prestigious Applicants of Intelligent Systems, PAIS 2004, Valencia, Spain, August 22-27, 2004*, pp. 858-862.

[Schlieder, 1996] Schlieder, C. (1996). Qualitative Shape Representation. *Geographic Objects with Indeterminate Boundaries*, pp. 123-140.

[Stock, 1997] Stock, O. (1997). *Spatial and Temporal Reasoning*. Kluwer Academic Publishers.

[Wolpert and Tickle, 2011] Wolpert, L. and Tickle, C. (2011). *Principles of Development*. OUP Oxford.

Qualitative Reasoning for Decision-Making: A Preliminary Report

Kenneth D. Forbus & Thomas Hinrichs
Qualitative Reasoning Group, Northwestern University
2133 Sheridan Road, Evanston, IL, 60201, USA
forbus@northwestern.edu

Abstract

Agents often have to make decisions with incomplete knowledge and few computational resources. We argue that qualitative representations and reasoning, especially combined with analogy, provide a natural approach to performing decision-making in situations with little data, incomplete models, and under tight computational constraints. Moreover, qualitative models provide a means of recognizing and framing decision problems. This paper describes our progress in exploring these ideas to date, using examples from experiments with a system that learns to play Freeciv, an open-source strategy game.

1 Introduction

Any agent operating in a world must make decisions. Many decisions are immediate, e.g. which way to turn during navigation or searching a physical environment. Reinforcement learning (Sutton & Barto, 2017) has often been used as a model for how to make such decisions. The tasks to be performed are fixed outside the learning mechanism, and the appropriate notion of state has been settled by the agent's designer (or evolution). Many other decisions require more analysis, such as how to optimize a supply chain, and the mathematical tools of decision theory are often brought to bear on these. Such tools require accurate mathematical models of the system to be optimized and accurate probabilities to handle uncertainty. Yet other kinds of decisions involve design: How should an efficient transportation network be built economically, or how much should be invested in what kinds of military units to provide an effective defense? These decisions involve organizational policies and values more directly, as well as having components of optimization, but with much less agreement on what quantitative models should be used, if any. What is common about all three kinds of decisions is that the formalisms used to address them do not incorporate the problem of framing the decision problems themselves. What should an agent pay attention to, what choices does it have to make, and how should it evaluate its progress? We believe that qualitative representations and reasoning can be used to capture aspects of decision-making that tend to be left implicit, just as they have been used to capture tacit knowledge in science and engineering. There has already been productive work on using qualitative representation in decision-support systems to be used by people (e.g. Agell et al. 2006; Benaroch & Dhar, 1995; Rosello et al. 2010;Zitek et al. 2009), but our focus here is on using QR within autonomous agents. The goal of reinforcement learning is to make good local decisions. The equivalent goal in qualitative decision-making is to make sensible decisions, and especially to avoid repeated blunders. We propose to do this by using analogy-based episodic memory, combined with qualitative representations, to detect problematic situations, build up models of their properties, and modify the agent's decision-making to avoid them in the future. The goal of traditional decision-theory is optimization. The equivalent goal in qualitative decision-making is to ensure that resource allocation is in alignment with the agent's priorites, as expresed by activations of goals. In keeping with our aim of formalizing more of the strategic thinking process itself, we use qualitative models to express strategies that are used to achieve goals (Hinrichs & Forbus, 2015). This includes synthesis goals, where setting up the means of production and handling investments is part of the problem.

This paper summarizes our work to date on using qualitative reasoning in decision-making. Most of this work has been done in a strategy game domain, Freeciv, so we start by briefly reviewing it. Then we discuss the use of qualitative representations to encode strategies, followed by a discussion of qualitative reasoning about resources. The problem of enabling agents to formulate their own evaluation metrics is discussed next, which leads naturally into how an agent can design systems in a domain that serve its strategic goals. Learning from experience via episodic memory is discussed next. We close with future work.

1.1 Freeciv in a nutshell

Freeciv[1] is a turn-based strategy game, played on a grid of tiles (see Figure 1). Players start with a few units – settlers can found cities, and workers can improve terrain, to make

[1] http://freeciv.org/

it more productive. Most of the map is unknown, and must be explored, by an explorer unit or a military unit. Unlike chess or go, new entities can be created by a player's cities. This includes new settlers, once the population grows enough, thereby fueling an expansion (and thus the need to find more terrain to build upon). Cities can be linked by

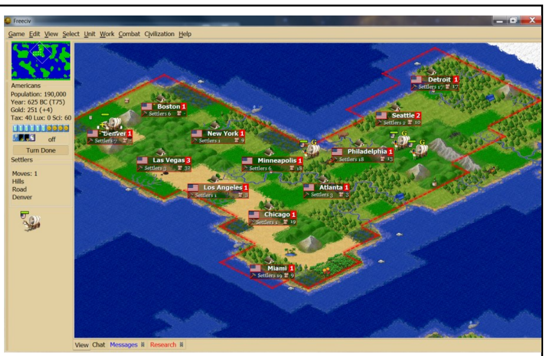


Figure 1: Freeciv

transportation networks, which again are constructed by the player using workers. Cities produce resources (food, production, luxury goods) that go into feeding citizens, expanding the population, and producing new units or buildings. Buildings enhance properties of a city: City walls improve its defensive capabilities, and a Library improves its science output. There are 40 types of city improvements and 51 types of units that can be built in Freeciv version 2.2.4, depending on which of 87 available technologies a player has achieved via research. Games are typically played on a 4,000 tile grid, and can last for hundreds of turns. Thus the sheer size contributes to the game's complexity.

There are several additional sources of complexity from the game's dynamics. There is finding good places to put cities, since some terrain has advantages, just like placing a city on a river or bay has advantages in real life. The kinds of units that can be built depends on what technologies the player has, which in turn depends on which technologies they choose to research. Some technologies directly allow the construction of new units: Catapults become possible once Mathematics is understood, for instance. Other technologies enhance properties of a civilization: Democracy, for example, enhances economic productivity but makes for citizen unrest if war is declared. Research investments trade off against production and food, leading to classic short-term spending versus long-term investment choices. And of course there are competing civilizations, with a simple diplomacy system and warfare, with units ranging from warriors to musketeers to nuclear weapons, depending on what technologies a civilization has gained (or stolen from others). Unlike some games, there can be serious disparities in technological advancement, based on a civilization's decisions: Archers trying to defend a city against tanks is a good lesson in the drawbacks of under-investing in research[2]. There are two ways to win: Either wipe out all other civilizations, or send the first starship to colonize Alpha Centari, which requires considerable research and economic prowess. Thus civilization-style games are extremely complex, much more so than chess or go.

Since Freeciv is open-source, and there is an active player community, quantaitive models are possible. Importantly, most players neither develop nor use them. Qualitative causal models of the game dynamics, combined with a sense of relative magnitudes and some spatial reasoning, suffice to play well in our experience. This makes it an excellent testbed for exploring qualitaitve reasoning in decision-making by autonomous agents.

Freeciv has been used by other AI researchers as well. Branavan et al. (2012) explored using Monte Carlo simulation and text analytics to construct a heuristic evaluation function. While it played well on a small subset of the game (smaller map, games ended at 75 turns), it required many trials to learn the game and used the game engine to do lookahead search while playing, a tactic which is not available for most domains. It also did not construct an inspectable model of causality in the domain, unlike our learned qualitative models. Ulam et al. (2008) investigated combining metareasoning and reinforcement learning for the subtask of city defense in Freeciv. While it uses model-based reasoning, the quantitative model it uses is constructed by hand, by contrast with our automatically learned qualitative models.

2 Strategic Planning as Qualitative Reasoning

We argue that continuous processes provide a representation for strategies (Hinrichs & Forbus, 2015). Consider the gap between a strategy, e.g. expand the cities in one's civilization, versus the actions actually available to carry out this strategy, e.g. build settlers, find terrain, send settlers out to establish new cities, and so on. An agent's individual actions are discrete, i.e. move one of it's units to a new tile (and thereby reveal the contents of adjacent squares, if not already revealed), build a city by using a settler (which is consumed in the process). Some actions are durative, e.g. irrigating a tile or building a road takes multiple turns. Formulating a crisp specific end goal would be very complex: A large continent can support a dozen cities and similar numbers of units. Exploration takes time, and so locations can only be planned as terrain becomes revealed. But in the meantime, other civilizations are building as well – there is a race for territory. So a strategy of doing a phase of data gathering followed by designing an optimal solution will be thwarted. Instead of this discrete, planning oriented model, we think instead of strategies as continuous processes that the agent implements by its choices of actions. Exploration is a process that increases the size of the pool of known tiles. Expansion is a continuous process that increases the size of the pool of a civilization's cities.

[2] We note that in Freeciv, research always succeeds and the benefits are accurately known in advance.

```
(isa Defending ModelFragmentType)
(genls Defending
      ProtectingSomething)
(participantType Defending
    protector-Agentive
    FreeCiv-MilitaryUnit)
(participantType Defending
    objectProtected FreeCiv-Actor)
(associatedRoleList Defending
   (TheList protector-Agentive
            objectProtected))
(participantConstraint Defending
  (and (objectFoundInLocation
         protector-Agentive
         objectProtected)
       (different protector-Agentive
                  objectProtected)))
(consequenceOf-TypeType Defending
  (qprop-
    ((QPQuantityFn Vulnerability)
         objectProtected)
    (DefensiveStrengthFn
      protector-Agentive
      FreeCiv-MilitaryUnit)))
```

Figure 2: Defense as a model fragment

Actions can be planned and evaluated based on whether they will ultimately contribute to implementing the processes that represent the agent's current strategy. This approach supports incrementality, an important property for dynamic worlds. The last few cities built, for instance, are typically created by settlers who were built in cities that did not even exist at the start of the game. Constructing extremely detailed long-range plans makes little sense in an adversarial situation, when units or terrain that are assumed turn out to no longer be available[3].

A concrete example will make this clearer. Consider the concept of defense. Defense isn't an action: Attacking an attacker is an action taken in the course of defending a city (or the unit itself), but is not the same thing. Defense isn't a state to be achieved, it is more about ensuing that an undesirable state (i.e. destruction or conquest) is prevented. Figure 2 illustrates a qualitaitve model of using a unit to defend a city or another unit. All concepts not in QP theory (Forbus, 1984) are from the OpenCyc ontology or our extensions of it for Freeciv. The key point is that the vulnerability of the object protected is reduced by the defensive strength of the protector. By expressing the defense of a civilization in terms of a sufficiently low vulnerability (discussed below), a limit point can be constructed for the process of adding defenses that adds defenses when the civilization becomes more vulnerable and stops building them when it estimates that it is sufficiently protected.

[3] As military commanders sometimes say, "The enemy has a vote."

3 Resources

A central concept in decision-making is the idea of resources. Some resources are the inputs to production or carrying out events: In Freeciv, there are several such resources. Gold provides a notion of money. Light bulbs (i.e. ideas) must be generated and accumulated to achieve a new technology. Shields are a unit of production, which is used in building new units or buildings in a city. Food is needed to keep a city alive, and when there is a surplus for a long enough period, the city's population grows. Cities produce these resources, based on where they are, how their citizens are put to work, tax rate settings, and what buildings have been created to improve a city. These trade off against each other, and the agent can exploit these tradeoffs in subtle ways. For example, cities along a hostile frontier might invest more in production, to create city walls and military units, while cities safely inside the civilization's borders might focus on research or economic advancement. These resources are also fungible: Gold can be spent to finish something a city is producing, e.g. city walls if there are barbarians approaching.

While these resources are represented in the game as integers, they can be effectively reasoned about as continuous quantities. Qualitative models describing the dynamics of such quanitites can be learned via demonstration, where an agent watches a human player (Hinrichs & Forbus, 2012) and by natural language instruction (McFate et al. 2014). These learning methods complement each other, and our agent uses a qualitative model of domain dynamics that combines knowledge learned by these methods. This model can be used to express overall goals of the game: Winning by military conquest is driving the number of enemy civilizations to zero, for instance.

One issue that arises with a complex set of goals is identifying tradeoffs. We automatically construct tradeoffs for type-level goals via a static analysis of the learned qualitative model (Hinrichs & Forbus, 2015). Goal tradeoffs can be characterized in terms of two dimensions: (1) Total versus Partial determines whether or not all instances of the goal must be adjusted in lockstep. For instance, how taxes are spent is determined at the level of the civilization, not individual cities, so that is a total tradeoff, whereas what is produced in cities can vary with the city. (2) Abrupt versus Progressive concerns whether the change in goals is instantaneous or can be gradually changed over time. Setting a tax rate is an abrupt action, whereas reducing emphasis on producing new settlers as a continent is filling is a progressive change in the relative priority of goals. These distinctions are independent, and hence there are different strategies for each of the four possible cases.

The nature of a *constructive* domain is that resources can be used to build new means of production (e.g. cities in Freeciv) or improve existing means (e.g. build a Factory in a city in Freeciv). While such resources are discrete, we find it useful to express goals about them in terms of continous properties. The cardinality of sets of some type of entity, such as number of cities, is a useful measure of progress in expansion. Sums across a civilization are another type of useful quantity for decision-making, e.g. overall research capacity, military strength, which can be defined by using a compositional sum (C+, from QP theory) over the appropriate types of entities (Hinrichs & Forbus, 2013).

Two other important resources that hold for almost any domain are space and time. In Freeciv, like today's planet, cities can only be built on land. For each new game, a new map is randomly generated. If a player is lucky enough to start on a large continent, they can focus on expansion and technologies for land-based units, leaving seafaring technologies for later, when their civilization is more advanced. If their continent is small (or even an island of a single tile), then their research priorities should instead focus on seafaring. Making this tradeoff requires taking information from exploration into account. Civilizations can span multiple continents, but this involves building sea units to transport other units (e.g. settlers, military units for protection) and coordinating such transportation. A landlocked civilization on a small continent is in a dismal place indeed, and may need to resort to warfare to expand. Such a strategy would involve first shifting production to military units, and then shifting back to settlers to grab new territory (and defend conquered cities). Thus the relative value of resources can shift drastically depending on the nature of the environment.

Time is perhaps the most subtle of resources. Adversarial domains often involve some sort of race, so the effective use of time becomes important. A qualitative model that stratifies durations of actions can be surprisingly useful in planning. Consider a city which is under threat by an enemy unit. It could switch production from what it is currently building, to create either city walls or a warrior (in the early game). Or a military unit can be moved from a neighboring city could be moved in to protect the city under threat. Switching production has a cost – which is moot if the city is conquored or destroyed, naturally – so avoiding that if possible would be good. Is the neighboring unit sufficiently close that it can make it in time? This depends in part on what transportion networks are available to the two units, their relative distance to the city, and how fast they can move (i.e. how many movement points per turn). In the early game, production is sufficiently slow that producing a defender in response to a perceived enemy threat is usually too late. This temporal consideration suggests a strategy of pre-positioning defenses and defenders before they are needed. Note that units and city improvements have upkeep costs, so this strategy (like all strategies) is not without drawbacks. Trying out alternate strategies, and keeping track of how well they succeed or fail, is a way of gathering data about the distribution of these relative intervals in a way that is directly relevant for decision-making. We plan to explore this by using analogical generalization (McLure et al. 2015), setting up a strategy, executing on it, and recording what happened afterwards, to learn which strategies work. One subtlety with dynamic worlds, of course, is that things change – in late-game strong civilizations, new technological advances may take only a turn or two, so researching a new technology and then building a needed unit based on that becomes a more viable strategy, whereas it is a recipe for defeat in the early game.

4 Formulating Evaluation Metrics

One of the key tasks of an agent in making decisions is deciding how to evaluate its alternatives. Rather than assuming a built-in evaluation function (as reinforcement learning does) or learning an evaluation for a single task (as inverse reinforcement learning does[4]), we believe that agents should formulate their own evaluation functions based on broad world knowledge as well as experience. An agent can be making many decisions at once, affecting a large set of ongoing strategies. Since one of the jobs of qualitative reasoning is framing problems, we view constructing evaluation metrics as one of the important tasks of QR for decision-making. We take *evaluation dimensions* to be parameters that can be approximated as continuous parameters. Every resource described above can be treated as an evaluation dimension, using either a continuous perspective on an integer quantity (e.g. gold, light bulbs) or integer quantities defined across sets (e.g. cardinality, totals). We denote the cost of an action or plan by the logical function `CostFn`. This function has two arguments: The plan itself and an evaluation dimension. Thus each evaluation dimension potentially provides a different way to look at the cost of an action or plan, and thereby enables tradeoffs to be explored. For instance, a plan ?p1 to reinforce a city under threat by moving a defender to it would have, as part of the constraints on any plan involving motion,

```
(qprop+ (CostFn ?p1 Time)
        (TravelTimeFn ?p1))
```

By contrast, a plan ?p2 to buy city walls in a city ?c would incur a cost in gold, which depends on how much effort had already been invested in building them:

```
(qprop+ (CostFn ?p2 Gold)
        (- (ProductionCostFn CityWalls
                             Shields)
           (ProductionSoFarFn CityWalls
                              ?c)))
```

The underlying game engine provides numerical values for some of these parameters (e.g. Shields) but not others (e.g. travel time). Such parameters are used by our systems to learn qualitative models from experimentation (e.g. Hinrichs & Forbus, 2007), but we do not use them for constructing

[4] We note that inverse reinforcement learning assumes that the expert traces it is observing are optimal – something that is not consistent with human decision making in complex domains (Kahaneman, 2011).

exact quantitative cost functions, because for constructive adversarial games in general, accurate mathematical models of the underlying domain are not available. Instead, we use experimentation to learn decision trees based on accumulating information from direct measurements. For example, a city built on grassland with wheat is much more productive than a city built on a desert. A learned decision tree that evaluates locations for city placement is used in the planning process for selecting expansion sites. Another method we plan to explore is using learned estimates of ordinal relationships to decide among alternatives. A rough estimate of the arrival time of the enemy would be enough to determine what alternatives (if any) are actually feasible.

5 Strategic Design

Some decisions are about how to build up new entities and systems to serve an agent's goals. In Freeciv, for example, building up a civilization entails creating a number of cities (as many as a dozen or more), improving the terrain around them, and linking them with roads (or railroads, once that technology is discovered). Such systems can be partially characterized by parameters whose settings should be learned by the agent, based on experience.

For example, how far apart should cities be? Claiming territory is useful, since it provides a buffer against enemies and gives an agent's civilization room to grow. On the other hand, unless each city has capable defenses, sending defenders to reinforce a city becomes more expensive. That suggests making the mean distance between cities smaller rather than larger. This consideration can be expressed qualitatively as follows: Consider ?p to be a generic plan involving travel between two cities, which can be approximated by the mean travel time:

```
(c+ (CostFn ?p Time) (TravelTimeFn ?p))
(qprop+ (TravelTimeFn ?p)
        (MeanCityDistFn ?civ))
```

implies

```
(qprop+ (CostFn ?p Time)
        (MeanCityDistFn ?civ))
```

On the other hand, trade routes are more valuable when two cities are far apart – if ?p is establishing a trade route between two cities, then

```
(qprop+ (ValueFn ?p Trade)
        (DistanceFn ?city1 ?city2))
```

Then taken across the entire civilization,

```
(qprop+ (ValueFn ?civ Trade)
        (MeanCityDistFn ?civ))
```

How do these qualitative models help in decision-making? They tell an agent about what relative likelihoods it needs to estimate. If warfare is likely to be common and trade is less important, keeping cities tightly clustered would be a better strategy. If bolstering trade is more important, then larger spacing might be a reasonable strategy. Building up models of what is likely in a game, via analogical generalization over episodic memories, could provide a way to estimate such likelihoods.

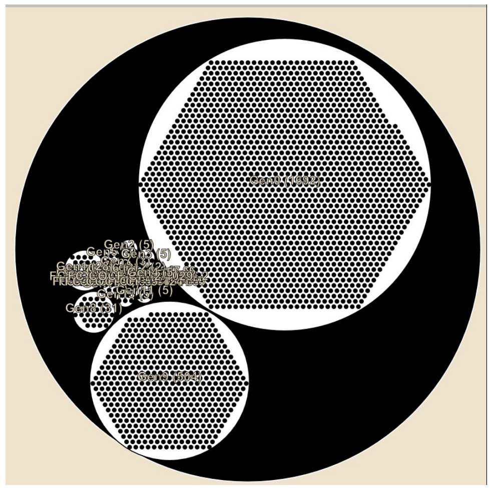


Figure 3: The generalization pool for doMove. The white circles are generalizations, the black dots within them represent particular experiences used in building that generalization.

6 Analogical Learning from Experience

Incomplete and incorrect models are the norm for agents operating in complex domains, especially when adversaries are involved, because other agents are often less predicable than domain physics. For example, a city can be weakened without being directly attacked by placing enemy units on the tiles immediately surrounding it, which prevents them from being worked and can cause starvation – an emergent behavior which is effectively a siege. A naïve agent can make suboptimal decisions, such as leaving military units out in the middle of nowhere, neither providing early warnings of approaching enemies nor defending anything. Trying to build settlers before a city has size 2 is impossible, because one citizen goes into the construction of a settler. All of these are things that human players figure out by watching their own behavior and learning how to improve it. We propose that analogical generaliztion over episodic memories provides a distillation of experience that can be used for such learning. This provides the rapid retrieval of either something to do, or something to avoid, a mechanism for the kind of human decision making that Kline describes in his recognition-primed decision model (Kline 1999).

For example, consider learning the immediate effects of actions. Most actions are fairly boring, either there is the same kind of change (i.e. changing production or research changes what a city is producing or the civilization is researching) or nothing happens for a while, if it is a durative action (e.g. irrigation has no immediate effect except for the worker no longer being idle). Movement is typically similarly boring, with one exception: Entering a hut. Huts on tiles can lead to multiple outcomes – gold or a new technology might be found, the unit might be killed by the inhabitants, or a new unit or city might be added to the civilization that made contact. Figure 3 illustrates a SAGE generalization pool for the primitive action doMove. Generalization pools accumulate examples of a concept incrementally, merging them into generalizations when they are suffiicently similar. Here the largest two generalizations are the typical outcomes of movement, with the different outcomes of entering a hut corresponding to smaller generalizations. Since SAGE constructs probabilities for each of the statements in every generalization pool, based on experience, the agent can compile a table of probabilities for the outcome of entering a hut (Table 1). Such experience-based probabilities are very useful for decision-making: Entering a hut can be seen here to be a good idea, overall, although given the chance of the unit being wiped out, diverting a settler on its way to found a new city to enter a hut is probably unwise.

Outcome of entering a hut	P
Gold Found	0.38
Technology Found	0.23
Unit joins your civ	0.23
City joins your civ	0.08
Killed by Barbarians	0.08

Table 1: Probability of outcomes for entering a hut, as calculated from SAGE's summaries of experience

How should a system know to build such a table? We have formulated a metric for surprise based on novelty concerning an experienced concept. That is, given an example E of a command C, whose analogical model consists of gpool(C), we define the novelty of E with respect to gpool(C) as

$$1\text{-}NSIM_B(BestMapping(SME(E,MACFAC(E,gpool(C)))))$$

That is, the base-normalized similarity score of the best mapping for the closest item retrieved from the generalization pool. If there is a case in the generalization pool that is identical to E (isomorphic up to entity renaming), then the numerical similarity will be 1, and E will have zero novelty. If nothing is retrieved, the numerical similarity is taken to be zero, and hence the novelty of E would be at its maximum, 1.0.

Not all novelty matters. SAGE provides a natural definition for novelty, since that can be taken as the dual of the decision that a new example is close enough for assiilation. That is, every generalization pool has an *assimilation threshold* A_t that ranges from 0 to 1. To respect this threshold, if an example E would be assimilated under the current threshold, then the novelty will be zero. The other factor which must be taken into account is how much experience the system has with the concept. We incorporate this factor by taking the product of the novelty with the following rate equation:

$$1\text{-}e^{-n/r}$$

Where *n* is the number of examples that have been added to the generalization pool so far, and *r* is a rate parameter, controlling how fast this asymptotes to 1. In the case of the doMove action, each time a new kind of outcome occurs when a hut is entered it signals a surprise, which can enable a system to keep track of that subset of actions as interesting.

Immediate effects of actions are just one kind of experience that should be routinely stored for subsequent analysis by an agent. Building up a model of the time that durative actions take can be done by taking before/after snapshots of the locale where such an action is taking place, and including the duration as part of the episodic memory. Statistics over those durations can then provide a robust way of estimating time costs for actions. In general, when decisions are made about an aspect of a domain that is not well understood, constructing episodic memories that capture what happened and how successful it was can be useful (e.g. worker assignments in Hinrichs & Forbus, 2007).

We see two other important functions of episodic memory. In adversarial domains, it is important to learn from what is being done to you, as well as what you do. Qualitative representaitons help lift descriptions to a level that is easier to compare, and hence to learn from. For example, the approach of an enemy unit to a city can be described as one interval using the Qualitative Trajectory Calculus (Van de Weghe et al. 2005) along with the duration of that activity, factoring out the specifics of the tiles traversed. Similarly, recognizing that a unit was lost because it was attacked by another unit is a very simple form of persective-taking. The other function of episodic memory is helping to set strategic parameters, e.g. what should the relative priorities of goals be, and what should limit points for strategic processes be? This, we suspect, is best done via a retrospective analysis of longer periods of play, abstracting out the specific events into statistics about global properties. For example, if a game was lost because an agent's cities were wiped out, then one potential solution is to increase the sensitivity to vulnerability, so that it pre-positions more defense resources and is more careful in future games.

7 Conclusions and Future Work

We believe that qualitative representations and reasoning can provide valuable services in formalizing the decision-making of agents in complex, dynamic adversarial worlds. The techniques outlined here complement traditional decision theory and reinforcement learning, since they are concerned with framing and formulating decision problems and using qualitative, causal models for both understanding the broad properties of domain dynamics and to express strategic concepts.

We plan to continue exploring these ideas in several ways. First, we plan to implement the other forms of episodic memory as outlined above, and explore their properties. Second, we plan to implement a reasoner that can formulate and articulate decision problems, criteria, and alternatives in a domain, so that agents can participate in joint problem solving involving strategic problems and learn more from natural langauge instruction, beyond the advice and domain-level causal models we have used language-based instruction for previously. Finally, we plan to explore cross-domain transfer: Tell a Companion stories about our world and ask what they imply about its strategies, and vice versa. Understanding when things will work in both (e.g. blockades) and when they won't (e.g. airlifts work in our world but not in Freeciv) is an important test of strategic thinking and transfer.

Acknowledgments

This research was supported by the Air Force Office of Scientific Research.

References

[Agell et al. 2006] Agell, N., Rovira, X., Sanchez, M., and Prats, F. (2006). Using Orders of Magnitude in Multi-Attribute Decision-Making. *Proceedings of QR 2006*.

[Benaroch & Dhar, 1995] Benaroch, M., & Dhar, V. (1995). Controlling the complexity of investment decisions using qualitative reasoning techniques. *Decision Support Systems*, 15(2):115-131.

[Branavan et al. 2012] Branavan, S., Silver, D., & Barzilay, R. (2012). Learning to win by reading manuals in a Monte-Carlo framework. *Journal of Artificial Intelligence Research*, 43:661-704.

[Forbus, 1984] Forbus, K. (1984). Qualitative process theory. *Artificial Intelligence*, 24, 85-168.

[Forbus & Hinrichs, 2018] Forbus, K.D. & Hinrichs, T. (2018). Analogy and Relational Representations in the Companion Cognitive Architecture. *AI Magazine*.

[Hinrichs & Forbus, 2012] Hinrichs, T. & Forbus, K. (2012). Learning Qualitative Models by Demonstration, *Proceedings of AAAI 2012*, pp.207-213.

[Hinrichs & Forbus, 2013] Hinrichs, T., & Forbus, K. (2013). Beyond the Rational Player: Amortizing Type-Level Goal Hierarchies.

[Hinrichs & Forbus, 2015] Hinrichs, T., and Forbus, K. (2015). Qualitative models for strategic planning. *Proceedings of the 3rd Annual Conference on Advances in Cognitive Systems*.

[Kahneman, 2011] Kahneman, D. (2011) *Thinking Fast and Slow*. Macmillan.

[Klein, 1999] Klein, G. (1999) *Sources of Power: How People Make Decisions*. MIT Press: Cambridge, MA.

[McFate et al. 2014] McFate, C., Forbus, K., & Hinrichs, T. (2014). Using Narrative Function to Extract Qualitative Information from Natural Language Texts. *Proceedings of AAAI 2014*.

[McLure et al. 2015] McLure, M., Friedman S. and Forbus, K. (2015). Extending Analogical Generalization with Near-Misses. *Proceedings of AAAI 2015*.

[Rosello et al. 2010] Rosello, L., Prats, F., Agell, N., & Sanchez, M. (2010). Measuring consensus in group decisions by means of qualitative reasoning. *International Journal of Approximate Reasoning*, 51:441-452.

[Sutton & Barto, 2017] Sutton, R. & Barto, A. (2017) *Reinforcement Learning: An Introduction*. 2nd Edition, MIT Press.

[Ulam et al. 2008] Ulam, P., Jones, J., & Goel, A. (2008) Combining Model-Based Meta-Reasoning and Reinforcement Learning for Adapting Game Playing Agents. *Proceedings of AIIDE 2008*.

[Van de Weghe et al. 2005] Van de Weghe, N., Cohn, A., De Tre, B., & De Maeyer, P. (2005) A Qualitative Trajectory Calculus as a basis for representing moving objects in Geographical Information Systems. *Control and Cybernetics*, 35(1):97-120.

[Zitek et al. 2009] Zitek, A., Schmutz, S., Preis, S., Salles, P., Bredeweg, B, & Muhar, S. (2009). Evaluating the potential of qualitative reasoning models to contribute to sustainable catchment management. *Ecological Informatics*, 4:381-395.

Qualitative Reasoning with Story-Based Motion Representations: Inverse and Composition

Juan Purcalla Arrufi*, Alexandra Kirsch†
*Human Computer Interaction, Universität Tübingen, Sand 14, 72076 Tübingen, Germany
juan.purcalla-arrufi@uni-tuebingen.de
† Independent Scientist, www.alexkirsch.de

Abstract

Representations of motion that are story-based constitute a promising tool to categorise the motion of entities, because they can be generated using any qualitative spatial representation, and they consider explicitly the speed of the entities. Up to the present, mainly categorisation properties of the story-based representations have been presented. In this paper we show how story-based representations allow for the reasoning operations that the qualitative calculi possess, namely, inverse and composition—We provide a method to compute the inverse, and a method that notably simplifies the computation of the composition.

1 Introduction and Related Work

Qualitative representations provide a way to transform and represent quantitative into qualitative knowledge [Hernandez, 1994; Dylla *et al.*, 2013; Dylla *et al.*, 2016]. In addition, they provide instruments for "reasoning" (in the broad sense of the term): the *'conceptual neighbourhood diagrams'* [Freksa, 1992], which enable decision-making [Dylla *et al.*, 2007]); also, operations between qualitative relations, such as *'inverse'* (also called *'converse'*), and *'composition'* are the base for reasoning methods—mostly constraint based techniques—in qualitative reasoning [Renz and Nebel, 2007; Cohn and Renz, 2008; Ligozat, 2012, Intr.].

To show that a qualitative representation is suitable for qualitatively representing quantitative knowledge, and to present its conceptual neighbourhood diagrams, e.g., [Van de Weghe and Philippe De Maeyer, 2005], is a more straightforward task than to show its suitability for reasoning in constraint based techniques, e.g., [Van De Weghe *et al.*, 2005a]—some qualitative representations remain, so far, without such a reasoning apparatus, e.g., [Glez-Cabrera *et al.*, 2013; Wu *et al.*, 2014], even when steps in this direction were taken [González-Cabrera *et al.*, 2010]. A cause is that the *composition* can only be computed by using the semantics of the relations [Renz and Nebel, 2007], and this often requires a burdensome manual case analysis, e.g., [Van de Weghe *et al.*, 2005b; Mossakowski and Moratz, 2010].

In previous work a novel method, the *'story-based'* approach, was presented to generate qualitative representations of motion [Purcalla Arrufi and Kirsch, 2017; Purcalla Arrufi and Kirsch, 2018] from any qualitative spatial representation. As example, the method was used to generate two representations of motion: Motion-RCC and Motion-OPRA$_1$, which are suitable to represent qualitative knowledge and perform decision-making. In this paper, we examine the possibility that such story-based representations may be suitable for qualitative reasoning: we provide a method for computing the inverse of relations and a method that notably simplifies the manual computation of the composition.

The paper is structured as follows: firstly, we introduce the story-based representations of motion by means of two examples. Secondly, we show how to compute the inverse of the story-based motion relations. Thirdly, we introduce our approximation method—an upper bound—to the composition set of *motion* relations; we call it *'narrative composition'* of motion relations; this method is based on another definition of composition between *spatial* relations. Further, the narrative composition of motion relations is also practically performed using paths on a matrix that we call *'composition matrix'*. Finally, we apply the method to concrete composition examples to show its validity.

2 Story-Based Representations of Motion

In this Section, we aim at categorising *'motion scenarios'* (Figure 1): pairs of two entities, k and l, that are described by two instantaneous pairs position-velocity vectors—one pair for each entity: $(\vec{x}_k, \vec{v}_k; \vec{x}_l, \vec{v}_l)$.

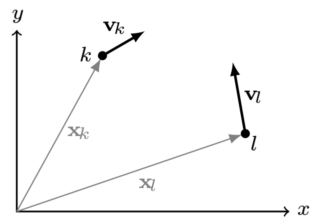


Figure 1: A *motion scenario*—two entities k and l, described by their instantaneous positions, $\vec{x}_k$ and $\vec{x}_l$, and velocities, $\vec{v}_k$ and $\vec{v}_l$.

Story-based representations of motion provide a categorisation for motion scenarios, moreover, these representations can be generated from any qualitative spatial representation $\mathcal{R}$.

In a story-based representation a relation is represented as $S_i(R_j)$; it relates the pair of entities (k, l). We call R_j the *'position component'*, and it is the qualitative spatial relation between (k, l) that is given by the generating spatial representation, $\mathcal{R}$. We call S_i the *'story component'* or just *'story'*: This is the full temporal (past; present; future) sequence of spatial relations, $(R_{i_1}, R_{i_2}, \ldots, R_j, \ldots, R_{i_m})$, that originates from a motion scenario by *assuming* uniform motion [Purcalla Arrufi and Kirsch, 2017] (See examples in Figures 2 and 4).

The total number of possible stories generated by a certain spatial representation is finite [Purcalla Arrufi and Kirsch, 2017, Math. App.], we call it the *'stories set'*. $\Sigma = \{S_1, S_2, \ldots, S_i, \ldots, S_n\}$.

2.1 Motion-RCC

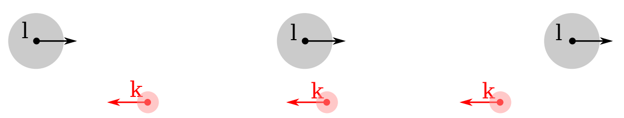
Motion-RCC [Purcalla Arrufi and Kirsch, 2018] is a story-based representation that is generated by the 8 spatial relations in the RCC representation [Randell *et al.*, 1992b]: DC, EC, PO, TPP, NTPP, EQ, TPPI, NTPPI; they are concerned with the overlapping of entities.

Motion-RCC has 16 stories, which, for the sake of simplicity, we reduce to five in this paper, by requiring that both entities, k and l, move with different velocities, and that the first entity, k, is smaller than the second one, l: {(DC), (DC, EC, DC), (DC, EC, PO, EC, DC), (DC, EC, PO, TPP, PO, EC, DC), (DC, EC, PO, TPP, NTPP, TPP, PO, EC, DC)}—we name this stories $\{S_{11}, S_{12}, S_{13}, S_{14}, S_{15}\}$.

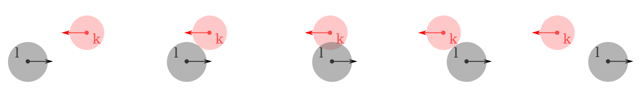
A Motion-RCC relation, e.g., $S_{13}(\mathrm{PO})$, is consequently formed by a story, S_{13}, and one of its spatial relations, PO, the position component; as we see in the third motion scenario in Figure 2b. There is a caveat: if the qualitative relation repeats in the story—as DC in the story S_{13}—we use the sub-index '$-$' or '$+$' to denote the temporal precedence: the relation $S_{13}(\mathrm{DC}_-)$ occurs previous to $S_{13}(\mathrm{DC}_+)$ (See first and last motion scenario in Figure 2b).

2.2 Motion-OPRA$_1$

The Motion-OPRA$_1$[Purcalla Arrufi and Kirsch, 2018] is generated by the OPRA$_1$ relations, i.e., $\{\angle^b_a \mid a, b \in \{0, 1, 2, 3\}\}$, which are concerned with orientations of entities (See Figure 3) [Moratz, 2006]. In Table 1 we detail the Motion-OPRA$_1$ stories that are used in the examples throughout the paper, for example, the story $S_{\mathrm{C}21}$ (Figure 4), used in the examples of Section 5˙.



(a) Example of story $S_{11} = (\mathrm{DC})$. Each motion scenario have motion relation $S_{11}(\mathrm{DC})$



$S_{13}(\mathrm{DC}_-)$ $S_{13}(\mathrm{EC}_-)$ $S_{13}(\mathrm{PO})$ $S_{13}(\mathrm{EC}_+)$ $S_{13}(\mathrm{DC}_+)$

(b) Example of story $S_{13} = (\mathrm{DC}, \mathrm{EC}, \mathrm{PO}, \mathrm{EC}, \mathrm{DC})$ where each pictured motion scenario has a particular motion relation.

Figure 2: Stories generated by the RCC spatial representation. All the spatial relations forming the story are represented by the corresponding motion scenario.

3 Computing Inverse Relations

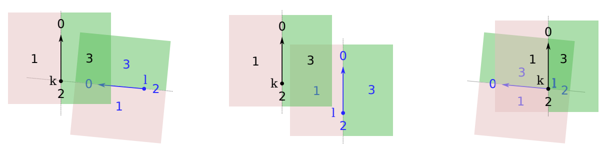
As we have explained in Section 2, each story-based relation is a binary relation between two entities, k and l, that has the form $S_i(R_j)$. S_i, is the story to which the motion scenario belongs, and R_j, the spatial relation of the motion scenario.

As $S_i(R_j)$ relates the pair (k, l), we solve a pertinent question: which is the story-based relation for the permuted pair, (l, k). This relation, $S_i(R_j)^{-1}$, is named *'inverse'* relation (or *'converse'*).

By definition, the terms S_i and R_j stand independently, as a Cartesian product, in the relations notation. Thus, $S_i(R_j)^{-1} = S_i^{-1}(R_j^{-1})$, and we can compute the inverse of each term just using the inverse of the generating spatial relation:

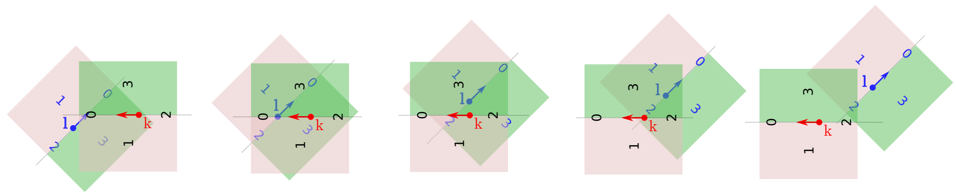
- R_j^{-1} is provided by the generating spatial representation. For example, $\mathrm{DC}^{-1} = \mathrm{DC}$, $\mathrm{TPP}^{-1} = \mathrm{TPPI}$, $(\angle^3_1)^{-1} = \angle^1_3$, or $(\angle 2)^{-1} = \angle 2$
- S_i^{-1} is computed by expressing the story as the list of spatial relations, then applying the inverse to any spatial relation on the list, and, finally, expressing the list of spatial relations as the corresponding story.
  For example, in Motion-RCC, the story S_{13}, is defined by the temporal sequence $(\mathrm{DC}, \mathrm{EC}, \mathrm{PO}, \mathrm{EC}, \mathrm{DC})$, thus $S_{13}^{-1} = (\mathrm{DC}^{-1}, \mathrm{EC}^{-1}, \mathrm{PO}^{-1}, \mathrm{EC}^{-1}, \mathrm{DC}^{-1}) = (\mathrm{DC}, \mathrm{EC}, \mathrm{PO}, \mathrm{EC}, \mathrm{DC})$, which is the same story S_{13}. Consequently, $S_{13}^{-1} = S_{13}$
  In Motion-OPRA$_1$ the story $S_{\mathrm{C}10}$ (See Table 1), corresponds to the temporal sequence $(\angle^3_1, \angle 3, \angle^1_3)$, thus the inverse is computed by the same procedure above: $S_{\mathrm{C}10}^{-1} = ((\angle^3_1)^{-1}, (\angle 3)^{-1}, (\angle^1_3))^{-1}) = (\angle^1_3, \angle 1, \angle^3_1) = S_{\mathrm{C}20}$

Thus, full examples of inverse relations in Motion-RCCare $S_{12}(\mathrm{EC}_+)^{-1} = S_{12}(\mathrm{EC}_+)$, $S_{13}(\mathrm{PO})^{-1} = S_{13}(\mathrm{PO})$. Examples in Motion-OPRA$_1$are $S_{\mathrm{C}10}(\angle^3_1)^{-1} = S_{\mathrm{C}20}(\angle^1_3)$, or, $S_{\mathrm{C}10}(\angle 3)^{-1} = S_{\mathrm{C}20}(\angle 1)$.



(a) $\angle^0_3$ (b) $\angle^1_3$ (c) $\angle 1$

Figure 3: Examples of OPRA$_1$ spatial relations.
$\angle^b_a$, between two entities k and l that are at different points. The syntax is $\angle^{k \text{ with respect to } l}_{l \text{ with respect to } k}$
$\angle a$, between two entities k and l that are at the same point. The syntax is $\angle_{\text{region of } k \text{ to which } l \text{ points}}$



$S_{C21}(\angle^1_3)$ $S_{C21}(\angle^1_0)$ $S_{C21}(\angle^1_1)$ $S_{C21}(\angle^2_1)$ $S_{C21}(\angle^3_1)$

Figure 4: Illustration of the story $S_{C21} = (\angle^1_3, \angle^1_0, \angle^1_1, \angle^2_1, \angle^3_1)$, with motion scenarios representing the full sequence of relations that constitute it.

4 Computing Composition of Story-Based Relations

The composition of qualitative relations is related to three entities k, l, and m. If we know the relation between the pair (k, l), say, R_A, and the relation between the pair (l, m), say, R_B, what would be then the possible relation for the pair (k, m)? Such relation or relations $\widetilde{R_C} = \{R_{C_1}, \ldots, R_{C_N}\}$ is called the *'composition'* of R_A and R_B—we express it symbolically as, $\widetilde{R_C} = R_A \circ R_B$. In the case of story-based relations we write $\widetilde{S_C(R_C)} = S_A(R_A) \circ S_B(R_B)$, i.e., $\widetilde{S_C(R_C)} = \{S_{C_1}(R_{C_1}), \ldots, S_{C_N}(R_{C_N})\}$

Usually, the result of the composition is not a single relation, but a set of relations. For example, in RCC, if the entities (k, l) fulfil the relation PO, and (l, m) fulfil TPP, then (k, m) may fulfil one of these relations: PO, TPP, or NTPP. We express that as, $\{\mathrm{PO}, \mathrm{TPP}, \mathrm{NTPP}\} = \mathrm{PO} \circ \mathrm{TPP}$. The full RCC composition values are displayed in [Randell *et al.*, 1992a], and an algorithm to find the composition in OPRA_1 is presented in [Mossakowski and Moratz, 2012].

Finding the composition is an arduous task that must be tailored to every representation. Though we do not exactly solve the composition for story-based relations, we take a step towards the computation of the composition by limiting its possible results. Indeed, we build an operation with story-based relations, the *'narrative composition'*, $S_A(R_A) \triangledown S_B(R_B)$, that yields a small superset of the standard composition, $S_A(R_A) \circ S_B(R_B) \subset S_A(R_A) \triangledown S_B(R_B)$, and usually does not coincide with it.

4.1 The Narrative Composition

In this Section, we describe how we build the narrative composition of story-based relations, $S_A(R_A) \triangledown S_B(R_B)$. We do it by requiring *necessary* conditions that the standard composition, i.e., $S_A(R_A) \circ S_B(R_B)$, must fulfil. Unfortunately, these conditions are *not sufficient*, i.e., the narrative compositions often adds extra relations to the standard composition (as shown in the example of Section 5.2). Nevertheless, the narrative composition gives a good first approach to the standard composition, that can be refined.

As a first step, we try to compute the story component of the composition, that is, we want to compute the stories S_C in the composition $\widetilde{S_C(R_C)} = S_A(R_A) \circ S_B(R_B)$. Stories are expressed as a finite sequence of spatial relations: $S_A = (R_{A_1}, R_{A_2}, \ldots, R_{A_i}, \ldots, R_{A_n})$ and $S_B = (R_{B_1}, R_{B_2}, \ldots, R_{B_j}, \ldots, R_{B_m})$. We notice that, by definition, the position component must be part of the corresponding story component, i.e., $R_A \in S_A$, and $R_B \in S_B$.

Any story $S_C \in \widetilde{S_C}$ must be necessarily related to some composition of the elements of S_A and S_B that keeps the temporal order. However, we show that such composition of elements is not directly given by the *standard* composition, i.e., $R_{A_i} \circ R_{B_j}$, but by a new type of composition of spatial relations, the *'narrative composition'*, which we express as $R_A \triangledown R_B$.

For example, if, in Motion-RCC, we want to compute the composition $\widetilde{S_C(R_C)} = S_A(R_A) \circ S_B(R_B)$, where $S_A = (\mathrm{DC}, \mathrm{EC}, \mathrm{DC})$ and $S_B = (\mathrm{DC})$, then, the pairwise standard composition of spatial relations produces this result: $(\mathrm{DC} \circ \mathrm{DC}, \mathrm{EC} \circ \mathrm{DC}, \mathrm{DC} \circ \mathrm{DC})$, which yields only *three-element* sequences: $(\mathrm{DC}, \mathrm{NTPPI}, \mathrm{PO})$, $(\mathrm{EC}, \mathrm{DC}, \mathrm{DC})$, and so on. This is unsatisfactory, because $S_C = (\mathrm{DC}, \mathrm{EC}, \mathrm{PO}, \mathrm{EC}, \mathrm{DC})$ is also a possible composition story that can never be obtained this way.

Summarising, we need to define new two operations: first, between spatial relations, $R_A \triangledown R_B$, second, between story-based motion relations, $S_A(R_A) \triangledown S_B(R_B)$. We name them both *'narrative composition'*—the operands, either motion or spatial relations, determine the way they are computed.

Narrative Composition of Spatial Relations: $\mathbf{R_A \triangledown R_B}$ The key to obtain all possible results, when composing stories, is that we have to use an *extended composition* between the elements, i.e., the spatial relations of the sequence. Indeed, in uniform motion, while (k, l) fulfil DC and (l, m) fulfil DC, it may happen that (k, m) goes through a sequence of relations, e.g., $(\mathrm{DC}, \mathrm{EC}, \mathrm{PO})$, where each element of the sequence is a possible result of the standard composition $\mathrm{DC} \circ \mathrm{DC}$.

Consequently, we define the *'narrative composition'*, $R_A \triangledown R_B$ of two spatial relations, R_A and R_B, as all possible subsequences of stories formed by relations that belong to the standard composition $R_A \circ R_B$.

Example 1: The standard composition of DC and EC yields five possible relations: $\mathrm{DC} \circ \mathrm{EC} = \{\mathrm{DC}, \mathrm{EC}, \mathrm{PO}, \mathrm{TPPI}, \mathrm{NTPPI}\}$. Consequently, $\mathrm{DC} \triangledown \mathrm{EC}$ is formed by all the combinations of relations in $\{\mathrm{DC}, \mathrm{EC}, \mathrm{PO}, \mathrm{TPPI}, \mathrm{NTPPI}\}$ that form story subsequences, i.e.,

$$
\textbf{Non-Parallel velocities} \quad \vec{v}_k \nparallel \vec{v}_l
$$

$$
\Sigma_C \left\{
\begin{array}{llll}
\multicolumn{4}{c}{\vec{v}_k \neq 0, \vec{v}_l \neq 0} \\
\angle^3_1\ \angle^0_1\ \angle^1_1\ \angle^1_2\ \angle^1_3 & S_{\text{C1-1}} & \angle^1_3\ \angle^0_3\ \angle^3_3\ \angle^3_2\ \angle^3_1 & S_{\text{C2-1}} \\
\angle^3_1\ \angle 3\ \angle^1_3 & S_{\text{C10}} & \angle^1_3\ \angle 1\ \angle^3_1 & S_{\text{C20}} \\
\angle^3_1\ \angle^3_0\ \angle^3_3\ \angle^2_3\ \angle^1_3 & S_{\text{C11}} & \angle^1_3\ \angle^1_0\ \angle^1_1\ \angle^2_1\ \angle^3_1 & S_{\text{C21}}
\end{array}
\right.
$$

$$
\Sigma_B \left\{
\begin{array}{llll}
\multicolumn{4}{c}{\vec{v}_k \neq 0, \vec{v}_l = 0} \\
\angle^3_1\ \angle^0_1\ \angle^1_1 & S_{B\text{1-1}} & \angle^1_3\ \angle^1_0\ \angle^1_1 & S_{B\text{2-1}} \\
\angle^3_0\ \angle 3\ \angle^1_2 & S_{B10} & \angle^0_3\ \angle 1\ \angle^2_1 & S_{B20} \\
\angle^3_3\ \angle^2_3\ \angle^1_3 & S_{B11} & \angle^3_3\ \angle^3_2\ \angle^3_1 & S_{B21} \\
\multicolumn{4}{c}{\vec{v}_k = 0, \vec{v}_l \neq 0} \\
\angle^1_3\ \angle^0_3\ \angle^3_3 & S_{B\text{3-1}} & \angle^3_1\ \angle^3_0\ \angle^3_3 & S_{B\text{4-1}} \\
\angle^1_0\ \angle 1\ \angle^3_2 & S_{B30} & \angle^0_1\ \angle 3\ \angle^2_3 & S_{B40} \\
\angle^1_1\ \angle^2_1\ \angle^3_1 & S_{B31} & \angle^1_1\ \angle^1_2\ \angle^1_3 & S_{B41}
\end{array}
\right.
$$

$$
\textbf{Parallel velocities} \quad \vec{v}_k \parallel \vec{v}_l
$$

$$
\Sigma_T \left\{
\begin{array}{ll}
\multicolumn{2}{c}{\text{The entities' trajectories are superposed}} \\
\angle^2_0\ \angle 0\ \angle^0_2 & S_{\text{T-1}} \\
\angle^0_0\ \angle 2\ \angle^2_2 & S_{\text{T0}} \\
\angle^0_2\ \angle 0\ \angle^2_0 & S_{\text{T1}}
\end{array}
\right.
$$

$$
\Sigma_P \left\{
\begin{array}{llll}
\multicolumn{4}{c}{\text{The entities' trajectories are not superposed}} \\
\multicolumn{4}{c}{\textit{singleton stories}} \\
\angle^3_1 & S_{P11} & \angle^3_3 & S_{P21} \\
\angle^1_3 & S_{P12} & \angle^1_1 & S_{P22}
\end{array}
\right.
$$

Table 1: Part of the OPRA$_1$ stories set, Σ, divided into meaningful subsets of stories: Σ_C, Σ_B, Σ_T, Σ_P

$\mathrm{DC} \triangledown \mathrm{EC} = \{(\mathrm{DC}), (\mathrm{DC}, \mathrm{EC}, \mathrm{DC}), \ldots, (\mathrm{EC}, \mathrm{PO}, \mathrm{EC}), \ldots, (\mathrm{EC}, \mathrm{PO}, \mathrm{TPPI}, \mathrm{PO}), \ldots\}$. Indeed (DC) is, for example, subsequence of story S_{12} (See section 2.1); (EC, PO, EC) is subsequence of S_{13}; and so on. However, a sequence such as $(\mathrm{EC}, \mathrm{TPPI}, \mathrm{EC})$, though it is a combination of relations from the composition $\mathrm{DC} \circ \mathrm{EC}$, does not belong to the narrative composition $\mathrm{DC} \triangledown \mathrm{EC}$, because it is not the subsequence of any story.

Narrative composition of Motion Relations: $\mathbf{S_A(R_A)} \triangledown \mathbf{S_B(R_B)}$ As a necessary condition, any story S_C belonging to the composition set $\widetilde{S_C(R_C)} = S_A(R_A) \circ S_B(R_B)$, must be the result of the narrative composition of the motion relations, i.e., $S_A(R_A) \triangledown S_B(R_B)$, which is obtained by concatenation of narrative composition of combinations of the spatial relations pairs constituting both stories S_A and S_B. Two further conditions complete the method: first, the composition of the position components, $R_A \triangledown R_B$, must always be present, because it corresponds to the position component of the composition, i.e., R_C; second, the concatenation of the narrative composition of spatial relations is only acceptable, if it is a story—not any sequence of concatenated narrative compositions is accepted.

The computation of the narrative composition is illustrated by means of Example 2, which is displayed in the composition matrix of Table 2.

Example 2: if we have two stories, $S_A = (R_{A_1}, R_{A_2}, R_{A_3})$, $S_B = (R_{B_1}, R_{B_2}, R_{B_3}, R_{B_4}, R_{B_5})$, the narrative story composition $S_A(R_{A_1}) \triangledown S_B(R_{B_3})$ is obtained by concatenating the narrative composition of story pairs of relations, keeping the temporal order, where the position components are also narratively composed, $R_{A_1} \triangledown R_{B_3}$ (grey boxes in Equations (1a) to (1c)). We present some possible results of narratively composed stories, which can be visualised as paths in a matrix (Table 2)

$$
\begin{aligned}
&(R_{A_1} \triangledown R_{B_1}, R_{A_1} \triangledown R_{B_2}, \boxed{R_{A_1} \triangledown R_{B_3}}, R_{A_1} \triangledown R_{B_4}, \\
&\qquad R_{A_1} \triangledown R_{B_5}, R_{A_2} \triangledown R_{B_5}, R_{A_3} \triangledown R_{B_5})
\end{aligned} \tag{1a}
$$

$$
\begin{aligned}
&(R_{A_1} \triangledown R_{B_1}, R_{A_1} \triangledown R_{B_2}, \boxed{R_{A_1} \triangledown R_{B_3}}, R_{A_2} \triangledown R_{B_4}, \\
&\qquad R_{A_3} \triangledown R_{B_5})
\end{aligned} \tag{1b}
$$

$$
\begin{aligned}
&(R_{A_1} \triangledown R_{B_1}, R_{A_1} \triangledown R_{B_2}, \boxed{R_{A_1} \triangledown R_{B_3}}, R_{A_2} \triangledown R_{B_3}, \\
&\qquad R_{A_2} \triangledown R_{B_4}, R_{A_3} \triangledown R_{B_4}, R_{A_3} \triangledown R_{B_5})
\end{aligned} \tag{1c}
$$

4.2 The Narrative Composition Matrix: Graphically computing the $\mathbf{S_A(R_A)} \triangledown \mathbf{S_B(R_B)}$

We have explained above how to compute the *narrative composition of motion relations*. In this Section we present the

$S_A(R_{A_1}) \triangledown S_B(R_{B_3})$	R_{B_1}	R_{B_2}	R_{B_3}	R_{B_4}	R_{B_5}
R_{A_1}	$R_{A_1} \triangledown R_{B_1}$	$R_{A_1} \triangledown R_{B_2}$	$R_{A_1} \triangledown R_{B_3}$	$R_{A_1} \triangledown R_{B_4}$	$R_{A_1} \triangledown R_{B_5}$
R_{A_2}	$R_{A_2} \triangledown R_{B_1}$	$R_{A_2} \triangledown R_{B_2}$	$R_{A_2} \triangledown R_{B_3}$	$R_{A_2} \triangledown R_{B_4}$	$R_{A_2} \triangledown R_{B_5}$
R_{A_3}	$R_{A_3} \triangledown R_{B_1}$	$R_{A_3} \triangledown R_{B_2}$	$R_{A_3} \triangledown R_{B_3}$	$R_{A_3} \triangledown R_{B_4}$	$R_{A_3} \triangledown R_{B_5}$

Table 2: *Narrative composition matrix* of the story-based motion relations $S_A(R_{A_1})$ and $S_B(R_{B_3})$. The corresponding stories are $S_A = (R_{A_1}, R_{A_2}, R_{A_3})$ and $S_B = (R_{B_1}, R_{B_2}, R_{B_3}, R_{B_4}, R_{B_5})$. The possible results of the narrative composition are obtained through paths in the matrix fulfilling the following properties: The yellow cells are start and end steps, the orange cell is a necessary middle step, because it is the narrative composition of the position components in $S_A(R_{A_1})$ and $S_B(R_{B_3})$. The three examples of valid paths (Equation (1)) are given, the *green* path corresponds to $(R_{A_1} \triangledown R_{B_1}, R_{A_1} \triangledown R_{B_2}, R_{A_1} \triangledown R_{B_3}, R_{A_1} \triangledown R_{B_4}, R_{A_1} \triangledown R_{B_5}, R_{A_2} \triangledown R_{B_5}, R_{A_3} \triangledown R_{B_5})$ (Equation (1a)); the *blue* path to $(R_{A_1} \triangledown R_{B_1}, R_{A_1} \triangledown R_{B_2}, R_{A_1} \triangledown R_{B_3}, R_{A_2} \triangledown R_{B_4}, R_{A_3} \triangledown R_{B_5})$ (Equation (1b)); and the *black* corresponds to $(R_{A_1} \triangledown R_{B_1}, R_{A_1} \triangledown R_{B_2}, R_{A_1} \triangledown R_{B_3}, R_{A_2} \triangledown R_{B_3}, R_{A_2} \triangledown R_{B_4}, R_{A_3} \triangledown R_{B_4}, R_{A_3} \triangledown R_{B_5})$ (Equation (1c))

$S_{12}(\mathrm{DC}_-) \triangledown S_{13}(\mathrm{PO})$	DC_-	EC_-	PO	EC_+	DC_+
DC_-	$\mathrm{DC}_- \triangledown \mathrm{DC}_-$	$\mathrm{DC}_- \triangledown \mathrm{EC}_-$	$\mathrm{DC}_- \triangledown \mathrm{PO}$	$\mathrm{DC}_- \triangledown \mathrm{EC}_+$	$\mathrm{DC}_- \triangledown \mathrm{DC}_+$
EC	$\mathrm{EC} \triangledown \mathrm{DC}_-$	$\mathrm{EC} \triangledown \mathrm{EC}_-$	$\mathrm{EC} \triangledown \mathrm{PO}$	$\mathrm{EC} \triangledown \mathrm{EC}_+$	$\mathrm{EC} \triangledown \mathrm{DC}_+$
DC_+	$\mathrm{DC}_+ \triangledown \mathrm{DC}_-$	$\mathrm{DC}_+ \triangledown \mathrm{EC}_-$	$\mathrm{DC}_+ \triangledown \mathrm{PO}$	$\mathrm{DC}_+ \triangledown \mathrm{EC}_+$	$\mathrm{DC}_+ \triangledown \mathrm{DC}_+$

Table 3: This table exemplifies a real Motion-RCC case of the general case in Table 2. We compose narratively the motion relations $S_{12}(\mathrm{DC}_-)$ and $S_{13}(\mathrm{PO})$; where the stories components are $S_{12} = (\mathrm{DC}_-, \mathrm{EC}, \mathrm{DC}_+)$, $S_{13} = (\mathrm{DC}_-, \mathrm{EC}_-, \mathrm{PO}, \mathrm{EC}_+, \mathrm{DC}_+)$. The difference are the *grey cells*—they correspond to *punctual relations*—therefore, we cannot directly step from grey cell into grey cell; this makes the black path an invalid path in this representation. However, the narrative compositions given by the blue and green are perfectly valid.

'narrative composition matrix', as a tool to perform the narrative composition of story-based relations of motion in a more visual way.

The narrative composition matrix of two motion relations, e.g., $S_A(R_{A_1}) \triangledown S_B(R_{B_3})$. is the table formed by the Cartesian product of all narrative compositions of the spatial relations that form the story components (Table 2).

Now we can see that the narrative composition of two motion relations are all possible stories defined by paths in the narrative composition matrix. The paths must fulfil the following conditions:

i) Every path begins in the upper left corner and ends in the lower right corner (yellow coloured cells in Tables 2 and 3)

ii) Every path can only be generated by moving from every cell either rightwards, downwards or diagonally rightwards downwards.

iii) Every path must pass through the cell containing the narrative composition of the position components, which is the spatial relation corresponding to the position component of the composition, i.e., R_C; (orange coloured cell in Tables 2 and 3). For example, in Table 2 is $R_{A_1} \triangledown R_{B_3}$.

4.3 Additional Constraints in the Composition Matrix

We can reduce the large number of possible paths, i.e., narrative compositions, if we have information about topological properties of the spatial relations.

In most spatial representations, amongst others RCC and OPRA, some relations can only occur at single time instant when they are part of a story. These *'punctual relations'* are

EC and TPP, in RCC, and every OPRA relation $\angle_x^y$ that contains 0 or 2 (e.g., $\angle_0^3$, or $\angle_2^2$).

Because every story is a continuous movement of entities in space (indeed, a uniform motion), we cannot transition directly between punctual relations, unless another relation is in between. As example, in Table 3, the punctual relations of Motion-RCC are marked in grey. Thus, the black path is no more a possible narrative composition of the stories, because it transitions from one grey cell into a another neighbouring grey one.

5 Examples of Narrative Composition in Story-Based Representations

In this Section we offer two full examples of *narrative* composition in the story-based relations of motion Motion-OPRA$_1$. Furthermore, we refine the results to obtain the *standard* composition between such relations.

5.1 Narrative Composition $\mathbf{S_{C21}}(\angle_{\mathbf{1}}^{\mathbf{3}}) \triangledown \mathbf{S_{T\text{-}1}}(\angle_{\mathbf{0}}^{\mathbf{2}})$

We want to compute the narrative composition $S_{\text{C21}}(\angle_1^3) \triangledown S_{\text{T-1}}(\angle_0^2)$ (See Figures 5 and 6)

Firstly, we express the stories as sequence of spatial relations: $S_{\text{C21}} = (\angle_3^1, \angle_0^1, \angle_1^1, \angle_1^2, \angle_1^3)$ and $S_{\text{T-1}} = (\angle_0^2, \angle 0, \angle_2^0)$. Secondly, we compute the narrative composition of stories by means of the narrative composition matrix (Table 4). The only possible path in the matrix that passes through the composed spatial components, i.e., $\angle_1^3 \triangledown \angle_0^2$ (blue cell), is the blue path.

The blue path in Matrix 4b generates many temporal sequences of relations, for example, $(\angle_0^1, \angle_1^1, \angle_2^1, \angle_1^2, \angle_0^3, \angle_1^3, \angle_3^3)$, and $(\angle_1^1, \angle_1^2, \angle_1^3)$. Only two of all generated sequences are Motion-OPRA$_1$ stories (see Table 1): S_{C21} and S_{B31}. The value of $\angle_1^3 \triangledown \angle_0^2$ compatible with the story yields the spatial component of the relation. Accordingly, $S_{\text{C21}}(\angle_1^3) \triangledown S_{\text{T-1}}(\angle_0^2) = \{S_{\text{C21}}(\angle_1^3), S_{\text{B31}}(\angle_1^3)\}$

However, looking at certain kinetic properties, we see that S_{B31} is not possible; because it implies that $\vec{v}_k = 0$, which contradicts $S_{\text{C21}}(\angle_1^3)$, where $\vec{v}_k \neq 0$. Therefore, the only possible result of the standard composition is the relation $S_{\text{C21}}(\angle_1^3)$ (Figure 7). Thus, finally, the standard composition yields $S_{\text{C21}}(\angle_1^3) \circ S_{\text{T-1}}(\angle_0^2) = S_{\text{C21}}(\angle_1^3) \subsetneq S_{\text{C21}}(\angle_1^3) \triangledown S_{\text{T-1}}(\angle_0^2) = \{S_{\text{C21}}(\angle_1^3), S_{\text{B31}}(\angle_1^3)\}$.

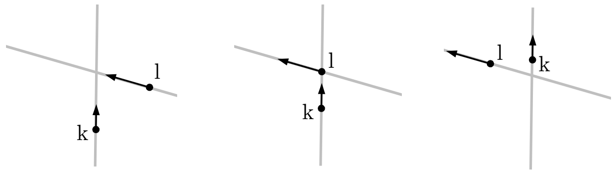
5.2 Narrative Composition $\mathbf{S_{C21}}(\angle_{\mathbf{3}}^{\mathbf{1}}) \triangledown \mathbf{S_{T\text{-}1}}(\angle_{\mathbf{2}}^{\mathbf{0}})$

Proceeding as in Section 5.1, we compute the narrative composition of relations using the narrative composition matrix (Table 4). In this case, the only possible path in the matrix that passes through the composed spatial components, i.e., $\angle_3^1 \triangledown \angle_2^0$ (red cell), is the red path.

If we analyse all possible sequences of relations that the red path generates by narrative composition, we obtain only five stories, which correspond to five motion relations: $S_{\text{C21}}(\angle_3^1) \triangledown S_{\text{T-1}}(\angle_2^0) = \{S_{\text{C2-1}}(\angle_3^1), S_{\text{C20}}(\angle_3^1), S_{\text{C21}}(\angle_3^1), S_{\text{B2-1}}(\angle_3^1), S_{\text{B3-1}}(\angle_3^1)\}$.

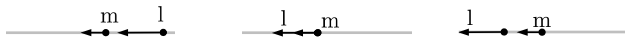
Analogous as in Section 5.1, we see that the relation $S_{\text{B3-1}}(\angle_3^1)$, is impossible—it implies that $\vec{v}_k = 0$, which is contradictory with one of the relations being composed, i.e., $S_{\text{C21}}(\angle_1^3)$, which implies that $\vec{v}_k \neq 0$. For the remaining compositions, we find possible realisations.

Therefore, the composition yields, $S_{\text{C21}}(\angle_3^1) \circ S_{\text{T-1}}(\angle_2^0) = \{S_{\text{C2-1}}(\angle_3^1), S_{\text{C20}}(\angle_3^1), S_{\text{C21}}(\angle_3^1), S_{\text{B2-1}}(\angle_3^1)\} \subsetneq S_{\text{C21}}(\angle_3^1) \triangledown S_{\text{T-1}}(\angle_2^0)$



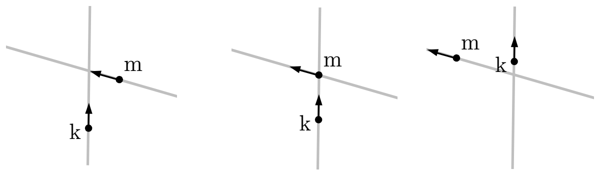
$S_{\text{C21}}(\angle_3^1)$ $S_{\text{C21}}(\angle_0^1)$ $S_{\text{C21}}(\angle_3^1)$

Figure 5: The story S_{C21}, displayed with some of its motion relations



$S_{\text{T-1}}(\angle_0^2)$ $S_{\text{T-1}}(\angle 0)$ $S_{\text{T-1}}(\angle_2^0)$

Figure 6: The story $S_{\text{T-1}}$ displayed with all motion relations involved



$S_{\text{C21}}(\angle_3^1)$ $S_{\text{C21}}(\angle_0^1)$ $S_{\text{C21}}(\angle_3^1)$

Figure 7: The solution of the composition $S_{\text{C21}}(\angle_1^3) \circ S_{\text{T-1}}(\angle_0^2)$ is the Motion-OPRA$_1$ relation $S_{\text{C21}}(\angle_1^3)$, belonging to the story S_{C21}, shown here.

6 Conclusion

We showed that the standard operations used for reasoning—inverse and composition—can be computed in the story-based representations of motion. Accordingly, such representations constitute a qualitative calculus, a more powerful tool than a categorisation. We gave a method, the narrative composition of story-based motion relations, to approximately compute the composition of motion relations—with this method we obtain a superset of the possible solutions that can notably approach the exact result. Therefore, the most appealing task for future work is to refine the narrative composition, so that we obtain a general method, if possible, to exactly compute the standard composition in any story-based motion representation.

Acknowledgements

Juan Purcalla Arrufi is grateful to Marcus Genkinger, CEO of STS Maschinenbau GmbH, for providing a well-paid job that can be flexibly adapted to his research activities. He is thankful to Martin Hinz and Jieen Chen for their prayerful support.

$S_{\text{C21}} \triangledown S_{\text{T-1}}$	$\angle^2_0$	$\angle 0$	$\angle^0_2$
$\angle^1_3$	$\angle^1_3 \triangledown \angle^2_0$	$\angle^1_3 \triangledown \angle 0$	$\angle^1_3 \triangledown \angle^0_2$
$\angle^1_0$	$\angle^1_0 \triangledown \angle^2_0$	$\angle^1_0 \triangledown \angle 0$	$\angle^1_0 \triangledown \angle^0_2$
$\angle^1_1$	$\angle^1_1 \triangledown \angle^2_0$	$\angle^1_1 \triangledown \angle 0$	$\angle^1_1 \triangledown \angle^0_2$
$\angle^2_1$	$\angle^2_1 \triangledown \angle^2_0$	$\angle^1_1 \triangledown \angle 0$	$\angle^1_1 \triangledown \angle^0_2$
$\angle^3_1$	$\angle^3_1 \triangledown \angle^2_0$	$\angle^3_1 \triangledown \angle 0$	$\angle^3_1 \triangledown \angle^0_2$

(a) *Narrative composition matrix* of the stories $S_{\text{C21}} = (\angle^1_3, \angle^1_0, \angle^1_1, \angle^2_1, \angle^3_1)$ and $S_{\text{T-1}} = (\angle^2_0, \angle 0, \angle^0_2)$.

$S_{\text{C21}} \triangledown S_{\text{T-1}}$	$\angle^2_0$	$\angle 0$	$\angle^0_2$
$\angle^1_3$	$\angle^1_{\{0,1,3\}}$	$\angle^1_3$	$\angle^1_{\{1,2,3\}}$
$\angle^1_0$	$\angle^1_1$	$\angle^1_0$	$\angle^1_3$
$\angle^1_1$	$\angle^1_{\{1,2,3\}}$	$\angle^1_1$	$\angle^1_{\{0,1,3\}}$
$\angle^2_1$	$\angle^2_1$	$\angle^2_1$	$\angle^{\{0,2\}}_{\{1,3\}}, \angle 1$
$\angle^3_1$	$\angle^3_{\{0,1,3\}}$	$\angle^3_1$	$\angle^3_{\{1,2,3\}}$

(b) We compute the composition in each cell of Matrix 4a (See the algorithm for OPRA$_1$in [Mossakowski and Moratz, 2012]).The narrative composition on each cell is the combination of the relations in the cell

Table 4: We compute the two examples in Section 5: $S_{\text{C21}}(\angle^3_1) \circ S_{\text{T-1}}(\angle^2_0)$, blue path passing through $\angle^3_1 \triangledown \angle^2_0$; and $S_{\text{C21}}(\angle^1_3) \circ S_{\text{T-1}}(\angle^0_2)$, red path passing through $\angle^1_3 \triangledown \angle^0_2$.
We can use the same narrative composition matrix, because the motion components of the relations, i.e., S_{C21} and $S_{\text{T-1}}$, are the same. The start and end cell for both examples are the same, they are yellow coloured. The cell of the current spatial component is coloured according to the path.

References

[Cohn and Renz, 2008] Anthony G Cohn and Jochen Renz. Qualitative Spatial Representation and Reasoning. In Vladimir Lifschitz van Harmelen and Bruce Porter, editors, *Handbook of Knowledge Representation*, volume 3 of *Foundations of Artificial Intelligence*, chapter 13, pages 551–596. Elsevier, 2008.

[Dylla *et al.*, 2007] F Dylla, L Frommberger, J O Wallgrün, D Wolter, B Nebel, and S Wölfl. SailAway: Formalizing navigation rules. In *Proceedings of the Artificial and Ambient Intelligence Symposium on Spatial Reasoning and Communication, AISB'07*, pages 1–5, 2007.

[Dylla *et al.*, 2013] Frank Dylla, Till Mossakowski, Thomas Schneider, and Diedrich Wolter. Algebraic Properties of Qualitative Spatio-temporal Calculi. In Thora Tenbrink, John Stell, Antony Galton, and Zena Wood, editors, *Spatial Information Theory*, volume 8116 of *Lecture Notes in Computer Science*, pages 516–536. Springer International Publishing, 2013.

[Dylla *et al.*, 2016] Frank Dylla, Jae Hee Lee, Till Mossakowski, Thomas Schneider, André van Delden, Jasper van de Ven, and Diedrich Wolter. A Survey of Qualitative Spatial and Temporal Calculi - Algebraic and Computational Properties. *CoRR*, abs/1606.0(212), 2016.

[Freksa, 1992] Christian Freksa. Temporal reasoning based on semi-intervals. *Artificial Intelligence*, 54:199–227, 1992.

[Glez-Cabrera *et al.*, 2013] Francisco J Glez-Cabrera, José Vicente Álvarez-Bravo, and Fernando Díaz. QRPC: A new qualitative model for representing motion patterns. *Expert Systems with Applications*, 40(11):4547–4561, 2013.

[González-Cabrera *et al.*, 2010] F J González-Cabrera, M Serrano-Montero, J C Peris-Broch, M T Escrig-Monferrer, and José Vicente Álvarez-Bravo. An Iconic Notation for describing the Composition between Relations of a Qualitative Model based on Trajectories. In *Distributed Computing and Artificial Intelligence (7th International Symposium)*, pages 657–664. Springer, 2010.

[Hernandez, 1994] D Hernandez. *Qualitative Representation of Spatial Knowledge*, volume 804 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin/Heidelberg, 1994.

[Ligozat, 2012] Gérard Ligozat. *Qualitative Spatial and Temporal Reasoning*. ISTE Ltd, London, 2012.

[Moratz, 2006] Reinhard Moratz. Representing relative direction as a binary relation of oriented points. In Gerhard Brewka, Silvia Coradeschi, Anna Perini, and Paolo Traverso, editors, *ECAI*, pages 407–411. 2006.

[Mossakowski and Moratz, 2010] Till Mossakowski and Reinhard Moratz. Qualitative Reasoning about Relative Direction on Adjustable Levels of Granularity. 2010.

[Mossakowski and Moratz, 2012] Till Mossakowski and Reinhard Moratz. Qualitative reasoning about relative direction of oriented points. *Artificial Intelligence*, 180-181:34–45, 2012.

[Purcalla Arrufi and Kirsch, 2017] Juan Purcalla Arrufi and Alexandra Kirsch. Using Stories to Create Qualitative Representations of Motion. In Ana Maria Olteteanu and Zoe Falomir, editors, *ProSocrates 2017, 2nd Symposium on Problem-solving, Creativity and Spatial Reasoning in Cognitive Systems*, number July, pages 19–28, Delmenhorst, Germany, 2017. CEUR Workshop Proceedings.

[Purcalla Arrufi and Kirsch, 2018] Juan Purcalla Arrufi and Alexandra Kirsch. Motion Categorisation: Representing Velocity Qualitatively. *Cognitive Systems Research*, 2018.

[Randell *et al.*, 1992a] D A Randell, A G Cohn, and Z Cui. Computing transitivity tables: A challenge for automated theorem provers. In Deepak Kapur, editor, *Automated Deduction—CADE-11*, pages 786–790, Berlin, Heidelberg, 1992. Springer Berlin Heidelberg.

[Randell *et al.*, 1992b] David a. Randell, Zhan Cui, and Anthony G Cohn. A Spatial Logic based on Regions and Connection. In *Third International Conference on Principles of Knowledge Representation and Reasoning (KR1992)*, pages 165–176, 1992.

[Renz and Nebel, 2007] Jochen Renz and Bernhard Nebel. Qualitative Spatial Reasoning Using Constraint Calculi. In Marco Aiello, Ian Pratt-Hartmann, and Johan Van Benthem, editors, *Handbook of Spatial Logics*, chapter 1, pages 161–215. Springer Netherlands, 2007.

[Van de Weghe and Philippe De Maeyer, 2005] Nico Van de Weghe and Philippe De Maeyer. Conceptual Neighbourhood Diagrams for Representing Moving Objects. *Lecture notes in computer science*, 3770:1–11, 2005.

[Van De Weghe *et al.*, 2005a] Nico Van De Weghe, Anthony G Cohn, Philippe De Maeyer, and Frank Witlox. Representing moving objects in computer-based expert systems: The overtake event example. *Expert Systems with Applications*, 29(4):977–983, 2005.

[Van de Weghe *et al.*, 2005b] Nico Van de Weghe, Bart Kuijpers, Peter Bogaert, and Philippe De Maeyer. A Qualitative Trajectory Calculus and the Composition of Its Relations. In M Andrea Rodríguez, Isabel Cruz, Sergei Levashkin, and Max J Egenhofer, editors, *GeoSpatial Semantics*, volume 3799 of *Lecture Notes in Computer Science*, pages 60–76, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[Wu *et al.*, 2014] Jing Wu, Christophe Claramunt, and Min Deng. Towards a Qualitative Representation of Movement. In Marta Indulska and Sandeep Purao, editors, *Advances in Conceptual Modeling*, volume 8823 of *Lecture Notes in Computer Science*, pages 191–200. Springer International Publishing, 2014.

Qualitative Environment Mapping Based on Unmanned Aerial Vehicles 3D Images

Adeline Sécolo[1], Paulo Santos[1], Zoran Sjanic[2], Piotr Rudol[2], Tommy Persson[2], Mariusz Wzorek[2], Patrick Doherty[2]
[1] Centro Universitário FEI, São Paulo, Brazil
[2] Linköping University, Linköping, Sweden
secolo@fei.edu.br, psantos@fei.edu.br, zoran.sjanic@liu.se, piotr.rudol@liu.se, tommy.persson@liu.se, mariusz.wzorek@liu.se, patrick.doherty@liu.se

Abstract

This paper describes a preliminary work towards a collaborative system involving humans and unmanned aerial vehicles to perform environment mapping using a qualitative description of 3D images. The images collected by a team of unmanned aerial vehicles flying over a region are preprocessed to reconstruct 3D images. The qualitative description of the 3D images aims to simplify the information delivered by the system to the users engaged in a team of search and rescue mission, enhancing their understanding and knowledge about the environment.

1 Introduction

The interaction between people and intelligent systems including robots, leads to a high demand of reliable representations of human knowledge. Remote-controlled robots are becoming a reality in the world, and its applications vary from delivery systems to complex surgeries. Specific aspects of interaction as cooperation and collaboration raises the ability of heterogeneous teams composed by humans and different kinds of robots, or with different resources, to solve difficult problems that rely on real time data analysis and quick responses.

Unmanned Aerial Vehicles (UAVs) are considered tools for emergency informatics [Murphy, 2016], which is a scientific field that approaches the use of different data-sets to save lives in natural or man-made disasters, through acquisition, organization and visualization of data and consequently which actions can be taken into account. Besides the use of UAVs for military proposes, there is potential for civil applications, as forest fire tracking, damage survey after earthquakes or tsunamis, monitoring of riots, search and rescue of missing people, and so on.

This paper presents an extension of the work developed in [Doherty *et al.*, 2016], where an integrated collaborative system (composed of humans and heterogeneous autonomous Unmanned Aerial Vehicles) pursues the goal of generating 3D models of an environment. The information contained in 3D images might not be readily interpreted by the final user, instead, a qualitative description that includes cognitive comprehensibility ensuring efficient understanding by humans without wide computational complexity can yield better results for tasks such as search and rescue missions.

In this work a set of subfields of Qualitative Spatial Reasoning are joined to provide a qualitative description of a scene, observed by humans and robots, from different aerial points of view, in a way that all agents involved in the system have the possibility to understand where the objects are and the relations between them. The reasoning presented in Region Connection Calculus, Allen's Interval algebra and Interval Occlusion Calculus are used to provide the possible relations between a pair of objects seen from an aerial point of view, while the Cardinal Direction Calculus provides the basis for the orientation definition of those points of view.

Three new relations, $above$, *in the top of* and *in the bottom of* are presented to describe Allen's relations $precedes$, $starts$ and $finishes$ seen from a lateral point of view. Combinations of the relations are also developed to describe the positions of the objects when the projections of their shapes in axles x and y are not totally aligned, and cannot be described by a single relation from Allen's set of relations. Humans and robots will be able to exchange information and infer data about their partial view of the domain.

2 Qualitative Spatial Reasoning

Representing and reasoning about spatial knowledge can naturally be done through description of relations between two or more objects, specifying how spatial entities are related in space with others, particularly when numerical information is unavailable or even unnecessary for humans [Chen *et al.*, 2015].

In this research a 3D space is considered and the spatial entities are defined by a set of spatial points and lines. The set of relations is JEPD (Joint Exhaustive and Pairwise Disjoint), considering that in the set of spatial relations there is one and only one that can be satisfied. JEPD relations can be called basic relations for representing definite relationships between spatial entities. The union of basic information can express indefinite information. All base relations are possible in case no information is known [Chen *et al.*, 2015].

2.1 Region Connection Calculus

The basic relations between spatial regions are described in the theory of *Region Connection Calculus (RCC)* [Randell *et al.*, 1992], through the primitive *Connected (C(x,y))*, where

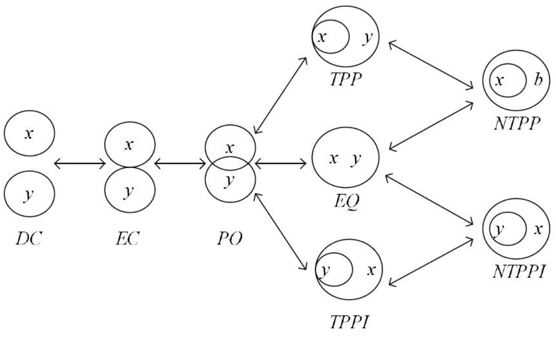


Figure 1: Conceptual neighborhood diagram
[Randell *et al.*, 1992]; [Renz, 2002].

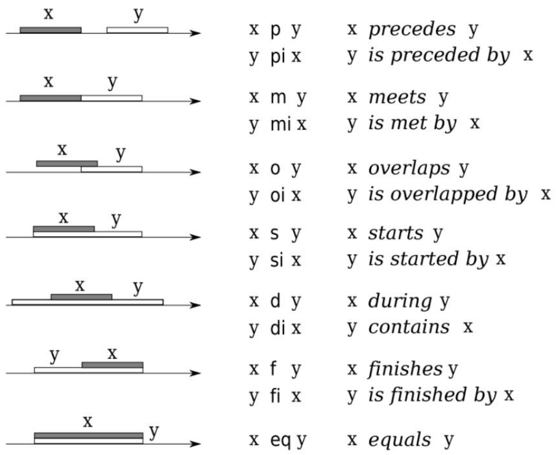


Figure 2: Allen's Relations
[Santos *et al.*, 2015], adapted from [Allen, 1990].

two non-empty regions of some topological space *x* and *y* are connected if and only if their topological closures share at least one common point.

Following the same theory, the constraint language RCC-8 contains eight JEPD base relations which allows reasoning about topological distinctions, with possibility to infer new spatial relations and transitions from incomplete spatial knowledge. Figure 1 shows the eight qualitative spatial relations covered by RCC-8 and its possible transitions illustrated by regions x and y on a conceptual neighborhood diagram [Randell *et al.*, 1992; Renz, 2002].

To define a reasoning for interpreting images collected by the UAVs, the RCC-8 theory is associated with Allen's Interval Algebra, described in the next section.

2.2 Allen's Interval Algebra

From the perspective of artificial intelligence, Allen's Interval Algebra describes a temporal representation and reasoning where temporal interval is considered as a primitive. This method represents the relationships between pairs of reference intervals taking into account their upper and lower limits in a hierarchical manner, resulting in a set of 13 jointly-exhaustive and pairwise-disjoint base relations [Allen, 1990]. Figure 2 shows Allen's relations considering two intervals, *x* and *y*.

Considering a constraint network pair *(N,C)*, *N* being the set of vertices where each domain element is represented by a vertex, and *C* representing the set of constraints defined by the basic Allen's relations, it is possible to verify the existence of a consistent scenario by imposing algebraic closure on the network of constraints, to confirm whether the configurations provided by the information from the agents are feasible in at least one scenario of a domain described by the set of Allen's relations [Santos *et al.*, 2015].

A network *(N,C)* is algebraically closed if its three vertices $(i, j, k) \in N^3$, and is consistent with the composition $C(i, j) \subseteq C(i,k) \circ C(k,j)$.

The Allen's relations can be applied to agents' viewpoints considering two rigid body convex entities, in an Euclidean plane. The $observers$ (also referred as $viewpoint$ Σ_i), are represented as pairs $\Sigma_i = (x_i, v_i)$, where x_i express the position of observer's centroid and v_i is the unit vector representing observer's orientation, and $objects$, identified by the 2D position of their centroids. The function $image(x, \Sigma)$ maps the projection of an object x perceived by an observer from a point of view Σ. In case more than one observer is present in the domain, each observer is perceived as an object when located within the field of view of another viewpoint [Santos *et al.*, 2015]. Furthermore, each observer is able to describe relations between pair of objects in its field of view.

An extension of Allen's Interval Algebra called *Interval Occlusion Calculus* was developed proposing to reason about occlusion from multiple points of view, and is described in the next section.

2.3 Interval Occlusion Calculus

Interval Occlusion Calculus (IOC) is a qualitative description of a set of basic relations between pairs of objects observed from a point of view, given the object's lines of sight [Santos *et al.*, 2015]. Considering that two objects A and B can be observed from a point of view Σ, the function $image$ defined by $a = image(A, \Sigma)$ and $b = image(B, \Sigma)$, maps the image of a physical body seen from a viewpoint Σ. Figure 3 illustrates an example of two bodies A and B, as well as a map based on the object's lines of sight, enabling the observers to locate themselves with respect to the qualitative relations observed between the images of the objects. If Σ is located on the region of the map marked by p, then the observer will notice that *a precedes b*, or if Σ is located on the region o^+, it could see that *a overlaps and is in front of b*. The same reasoning applies to all other positions of the map. Notations "region 1, 2, 3" and the red dashed region will be explained hereafter.

According to the distance between the observer and the objects (the depth), the image projected in a point of view Σ can be bigger or smaller. A layered interval is defined by $I = (I_a, l)$ where $I_a = (x_1, x_2)$, $x_1 < x_2$ are real numbers, x_1 is the lower limit of I, x_2 represents its upper limit and l is the $layer$ of I. The function $ext(I)$ maps the extension (the upper and lower limits) of a layered interval I and the function $l(I)$ maps the proximity between the object and the observer (as closer observers and objects are, greater is the

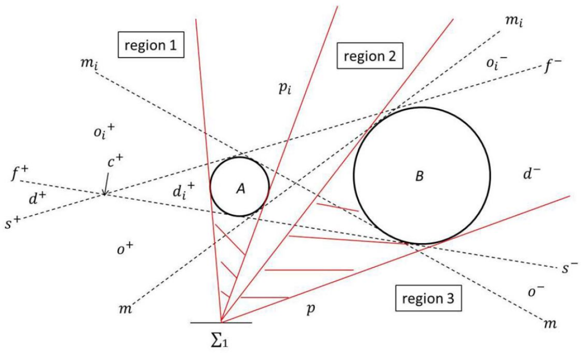


Figure 3: Basic Relations of Interval Occlusion Calculus [Santos *et al.*, 2015].

value of l). Finally, for any IOC relation r between two intervals I and J, the notation *I r* $^+$ *J* applies iff $l(I) > l(J)$ and *I r*$^-$ *J* applies iff $l(I) < l(J)$ [Santos *et al.*, 2015].

It is possible to apply IOC relations for any pair of objects, including environments where multiple agents are able to share each other's their spatial observations from different perspectives, interpreting and checking the consistency of the information received from other agents according to the translation table that can be found in [Santos *et al.*, 2015]. Let two distinct point of views Σ_1 and Σ_2 observe two objects A and B. The translation table gives the set of possible relations R considering exhaustively all the possible locations of Σ_2 from the point of view Σ_1, taking into account that *a = image (A,* Σ_1*), b = image (B,* Σ_1*), a' = image (A,* Σ_2*), b' = image (B,* Σ_2*),* σ_1 *= image (*Σ_1*,* Σ_2*)* and σ_2 *= image (*Σ_2*,* Σ_1*)*.

Reasoning about the environment showed in Figure 3, it is possible to interpret that if Σ_1 sees $\{a\ p\ b\}$, $\{\sigma_2\ p_i\ a\}$, and $\{\sigma_2\ p\ b\}$; then Σ_2 is in Region (2) of the map and its possible observations is the set of relations $\{a'\ \{p_i; p\}\ b'\}$. As another example, if Σ_2 is in the red-dashed region between A and Σ_1, then Σ_1 observes that $\{\sigma_2\ \{s^+, f^+, d^+\}\ a\}$, and the set of relations from Σ_2 would be $\{a'\ \{p, m, o^+\}\ b'\}$ [Santos *et al.*, 2015].

2.4 Cardinal Direction Calculus

A qualitative reasoning to deal with different perspectives perceived by multiple agents observing a scene is called Cardinal Direction Calculus.

The Cardinal Direction Calculus (CDC) is a formalism to reason about cardinal directions between objects and its relations. CDC is composed by 9 basic relations including a neutral region: *north (n), east (e), west (w), south (s), northwest (nw), northeast (ne), southeast (se), southwest (sw)* and *equal (eq)*, that corresponds to the neutral region [Frank, 1996].

Through CDC relations it is possible to infer the direction of two objects I and J from the knowledge of the direction between I and K, and between K and J.

In [Frank, 1996] it is possible to find more information related to cardinal reasoning, as a composition table of projection-based directions, a study about cone-shaped area for which a symbolic direction is applicable and the quadrant-projection that includes the neutral area.

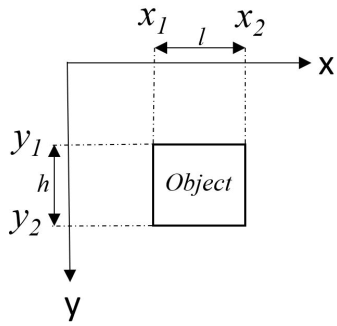


Figure 4: Object's projections from an aerial point of view.

The multiple point of views and relations mentioned in this chapter will be joined with the collaborative reasoning to be described in next sections in order to provide the appropriate constraints to interpret 3D images.

3 Collaborative Spatial Reasoning for Environment Mapping

Assuming that a system is composed by two or more agents and that one or more of those agents are humans, the mission to describe the position of objects present at a previously unknown environment might be impaired without the use of sentences that take into account the way people are used to describe this kind of information.

Considering an image composed by two objects I and J and observed from an aerial point of view Σ pointing to a direction v, the observed relations r are described in the sentence *I r J:* Σ, *v* and based on Allen's Interval Algebra, Interval Occlusion Calculus (IOC) and Region Connection Calculus, described on Section 2.

The image of the objects seen from an aerial point of view is processed by projecting its shape on two axes, x and y, whose origin is located in the upper left corner of the image (Figure 4). Whilst in IOC a single layered interval is considered, an aerial view requires at least two intervals: l and h where $l = (x_1, x_2)$; $x_1 < x_2$ and $h = (y_1, y_2)$; $y_1 < y_2$. To reason about 3D images, an additional interval c where $c = (z_1, z_2)$; $z_1 < z_2$ is added with the projections of the object on axis z. The formalization of this theory and a deep study about its implementation can be seen in [Guesgen, 1989] and [Balbiani *et al.*, 2002].

The cardinal direction indicated by a compass in the case of robotic agents and by human sense of direction in case of human agents, determines the vector v to represent the orientation of a point of view Σ observing the objects in a scene. The aerial view can be made by a camera on board the drones or a person looking at an environment from the window of a building, a range of $\pm$ 22,5º is considered to classify each cardinal direction, as showed on Figure 5. The logic to describe the observed relations for objects present in a scene requires the possibility to be transferred, to correctly describe relations from any point of view, for example, if object I precedes object J when Σ_1 is directed to $North$, the same scene seen from a Σ_2 directed to $South$ will be described as object J precedes object I. The directions v considered are: *North*

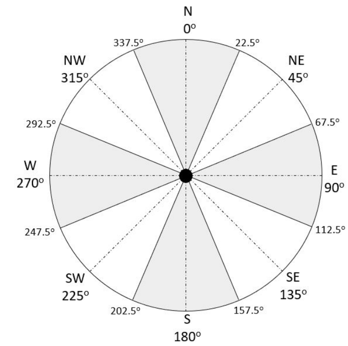


Figure 5: Cardinal directions.

(N) = 0^o, North East (NE) = 45^o, East (E) = 90^o, South East (SE) = 135^o, South (S) = 180^o, South West (SW) = 225^o, West (W) = 270^o and *North West (NW) = 315^o*.

The notation "*Object$_1$ relation Object$_2$: Σ, v*" is used to describe the possible relations between two objects observed from an aerial point of view Σ oriented to the direction v. Extending the single layered interval to a representation in two axes, more than one relation can apply to a pair of objects, resulting in the first set of reflexive relations described below, considering the arrangement of objects as shown in Figure 6. The relations written in capital letters refers to the RCC set of relations.

- *I p J : Σ, v*, read as "I precedes J from Σ oriented to v" and defined by *l(I) p l(J) $\cap$ l(I) DC l(J)*;
- *I $\{$p, m$\}$ J : Σ, v*, read as "I precedes and meets J from Σ oriented to v" and defined by *l(I) p l(J) $\cap$ l(I) EC l(J)*;
- *I $\{$p, o$\}$ J : Σ, v*, read as "I precedes and overlaps J from Σ oriented to v" and defined by *l(I) p l(J) $\cap$ l(I) O l(J)*;
- *I s J : Σ, v*, read as "I starts J from Σ oriented to v" and defined by *l(I) s l(J) $\cap$ h(I) TPP h(J)*;
- *I d J : Σ, v*, read as "I is during J from Σ oriented to v" and defined by *l(I) d l(J) $\cap$ h(I) NTPP h(J)*;
- *I f J : Σ, v*, read as "I finishes J from Σ oriented to v" and defined by *l(I) f l(J) $\cap$ h(I) TPP h(J)*;
- *I eq J : Σ, v*, read as "I is equal to J from Σ oriented to v" and defined by *l(I) eq l(J) $\cap$ h(I) eq h(J)*.

The same object's positions showed in Figure 6 seen from a lateral point of view results in the second set of relations of Figure 7. The relations $precedes$, $starts$ and $finishes$ are represented by three new relations: $above$, top and $bottom$, defined below, to avoid misunderstandings regarding object's positions when seen from different directions. The relation $during$ is not described here because it keeps the same definition from the first set of relations.

- *I a J : Σ, v*, read as "I is above J from Σ oriented to v" and defined by *h(I) a h(J) $\cap$ h(I) DC h(J)*;
- *I $\{$a, m$\}$ J : Σ, v*, read as "I is above and meets J from Σ oriented to v" and defined by *h(I) a h(J) $\cap$ h(I) EC h(J)*;

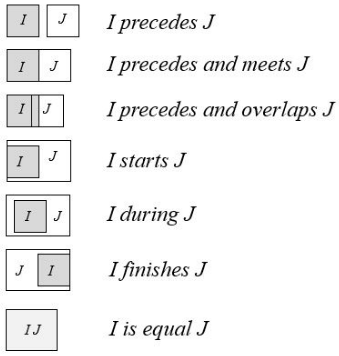


Figure 6: First set of relations and its definitions derived from Allen's Relations Calculus and Region Connection Calculus.

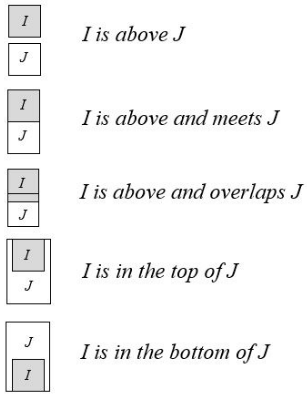


Figure 7: Second set of relations and its definitions derived from Allen's Relations Calculus and Region Connection Calculus.

- *I $\{$a, o$\}$ J : Σ, v*, read as "I is above and overlaps J from Σ oriented to v" and defined by *h(I) a h(J) $\cap$ h(I) O h(J)*;
- *I t J : Σ, v*, read as "I is in the top of J from Σ oriented to v" and defined by *h(I) t h(J) $\cap$ l(I) TPP l(J)*;
- *I b J : Σ, v*, read as "I is in the bottom of J from Σ oriented to v" and defined by *h(I) b h(J) $\cap$ l(I) TPP l(J)*.

A third set of relations was developed associating the first and second sets to express the relations for objects whose projections l and h are not totally aligned, and here the formalization is different from [Guesgen, 1989] because even working with intervals projected in two axles, it can express three possible relations, while in [Guesgen, 1989] it is possible to represent only two relations. Figure 8 shows the position of the objects described below.

- *I $\{$p, a$\}$ J : Σ, v*, read as "I precedes and is above J from Σ oriented to v" and defined by *l(I) p l(J) $\cap$ h(I) a h(J) $\cap$ l(I) DC l(J)*;
- *I $\{$p, m, a$\}$ J : Σ, v*, read as "I precedes, meets and is above J from Σ oriented to v" and defined by *l(I) p l(J) $\cap$ l(I) EC l(J) $\cap$ h(I) a h(J)*;
- *I $\{$p, o, a$\}$ J : Σ, v*, read as "I precedes, overlaps and is

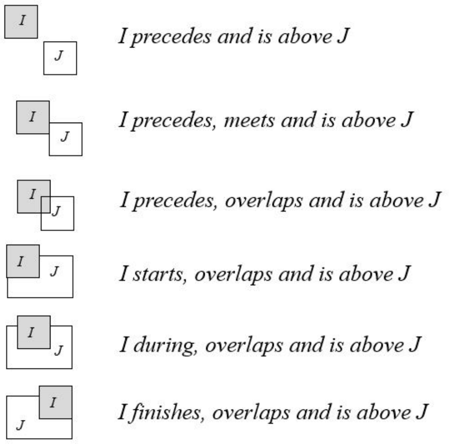


Figure 8: Third set of relations and its definitions derived from Allen's Relations Calculus and Region Connection Calculus.

above J from Σ oriented to v" and defined by *l(I) p l(J)* $\cap$ *l(I) PO l(J)* $\cap$ *h(I) PO h(J)* $\cap$ *h(I) a h(J)*;

- *I* $\{s, o, a\}$ *J :* Σ, v, read as "I starts, overlaps and is above J from Σ oriented to v" and defined by *l(I) s l(J)* $\cap$ *h(I) PO h(J)* $\cap$ *l(I) PO l(J)* $\cap$ *h(I) a h(J)*;
- *I* $\{d, o, a\}$ *J :* Σ, v, read as "I is during, overlaps and is above J from Σ oriented to v" and defined by *l(I) d l(J)* $\cap$ *h(I) PO h(J)* $\cap$ *l(I) PO l(J)* $\cap$ *h(I) a h(J)*;
- *I* $\{f, o, a\}$ *J :* Σ, v, read as "I finishes, overlaps and is above J from Σ oriented to v" and defined by *l(I) f l(J)* $\cap$ *h(I) PO h(J)* $\cup$ *l(I) PO l(J)* $\cap$ *h(I) a h(J)*.

To conclude, it is perceived that the basic relations *precedes, meets, overlaps, starts, during, finishes, equal* and the new ones *above, top* and *bottom* are enough to describe all possible configurations for a pair of objects.

Qualitative solutions for the robotic navigation problem was discussed in [Schlieder, 1993] and [Santos *et al.*, 2016], where reference points was encoded in the tessellation of a plane into regions to enable agents to navigate and localize themselves in an environment through the identification of landmarks. This approach takes the occlusion factor as an important role, making it difficult to implement the same reasoning for an aerial point of view, as the occlusion identified in most cases are in the vertical plane, for example, the leafs occluding a vehicle parked under a tree.

The effectiveness of the logic presented in this chapter will be tested using video image collected by drones in an experiment described in Section 5.

4 Collaborative Reasoning

Multi-agent systems are often supposed to have several advantages over single robot systems, as the capability to accomplish a single task faster, or efficiently exchange information about their position to precisely localize themselves whenever they sense each other [Burgard *et al.*, 2000]. When multiple agents are observing a scene and occlusion prevents sensors from assessing parts of the objects present in the scene, distinct observers can provide multiple view-point descriptions about the objects they can detect, improving the completeness of the information observed [Santos *et al.*, 2015].

Collaborative systems are required to integrate not only robots, but also human resources. For complex emergency scenarios such as search and rescue missions, providing assistance and guiding people to a safe destination from inhabitants lost in wilderness regions, at sea scenarios, places devastated by earthquakes, flooding or forest fires [Doherty *et al.*, 2010] it is an essential role to provide an effective communication and continuous interaction between people and robotic agents in order to achieve mission goals, specially the ones related to environment exploration, so that each individual can explore different areas simultaneously [Burgard *et al.*, 2000].

5 Experiment

Three data collection flights were performed at Motala Flygklubb and two different UAV platforms were used.

The first one is a DJI Matrice 600 Pro[1] research platform and The second platform is a DJI Matrice 100[2].

Both platforms are equipped with Intel NUC computers using Core i7-7567U processors, 16 GB of memory, and 500 GB SSD storage. The DJI Zenmuse Z3[3] cameras were used to collect video and images during the experimental flights.

The scanning patterns flown to collect the experimental data were automatically generated to cover a designated area. The platforms took off manually and after reaching a safe altitude flied autonomously over the region and performed the scan according to the pre-determinate altitude. The flights were performed at three different altitudes: 30, 50 and 80 m above the ground level (AGL).

The videos and photos collected register a partial view of the environment and were taken from different directions, to show different partial perspectives of each scene. The images contain many buildings with different sizes and colors, some cars, trees and a small airplane. The background contains areas covered by grass, asphalt and a road.

Each drone flew over a different portion of the area, but capturing images of some objects in common, in order to use those objects as references to infer the position of all other objects that are not in its field of view. The information can be easily exchanged by agents via wi-fi, including the ground operator. It is possible to transmit images, but a large broadband is necessary and it is not always available. The developed sentences in the format "$Object_1$ $\{relations\}$ $Object_2$: Σ, v" can be transmitted using only sparse data and thus cope with the bandwidth problem, independent of the number of axles used to generate the relations. The complexity will be related to the process of object identification, projections on axles and relations generation, and not to data transmission.

[1] `www.dji.com/matrice600/info`
[2] `www.dji.com/matrice100`
[3] `www.dji.com/zenmuse-z3`

Figure 9: Image captured by drone DJI Matrice 100, flying in the East direction

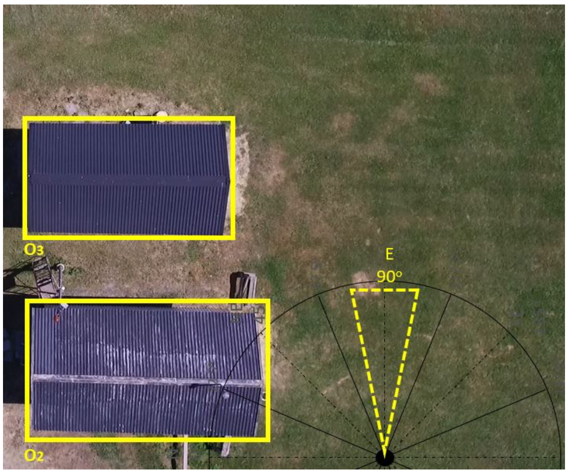


Figure 10: Image captured by drone DJI Matrice 100, flying in the East direction

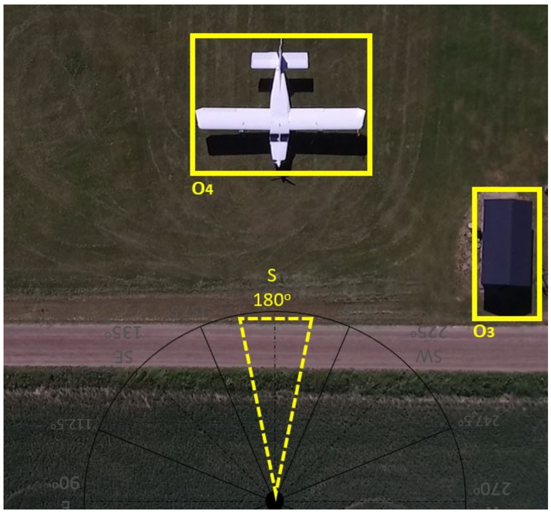


Figure 11: Image captured by drone DJI Matrice 600, flying in the South direction

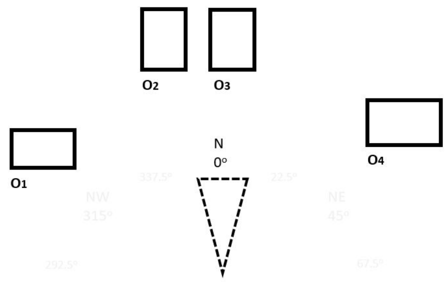


Figure 12: Sketch of a visual map representing object's positions

Figures 9 and 10 show images captured by one drone flying in the East direction, at an altitude of 30 meters. As soon as the system recognize a pair of objects, each object is labeled and the relations between them are identified according to the set of restrictions of Section 3, and the information is transmitted to the other agents. The information contained in Figure 9 is transmitted in the format: $Object_2\ \{p,\ a\}\ Object_1 : \Sigma_1,\ E$, while Figure 10 should transmit the information $Object_3\ \{a\}\ Object_2 : \Sigma_1,\ E$.

The second agent produced the image seen in Figure 11 flying in the South direction, at an altitude of 50 meters. The information transmitted by this agent is: $Object_4\ \{p,\ a\}\ Object_3 : \Sigma_2,\ S$. The airplane labeled as $Object_4$ is not identified as a square or rectangle in the image, but its extreme points can be projected in the axes enabling the system to fit it under a square format.

The information exchanged by the agents is supposed to produce a result similar to the Figure 12, where it is possible to see a sketch of a map with the objects translated to the North direction, to complement the information already received in the format of sentences. The Figure 13 shows the environment used in this example. The image was captured by DJI Matrice 600, flying to North direction, at 80 meters of altitude.

Figure 13: Real environment

Currently, those images and videos are under analyses to evaluate which computer vision techniques and filters will be applied to identify and label the objects present in each scene or frame for posterior implementation and test of the relations of Collaborative Spatial Reasoning.

6 Next Steps

After implementation, test and correction of any inconsistencies found in the Collaborative Spatial Reasoning for Environment Mapping using videos and 2D photographs captured by drones, the same reasoning will be applied to 3D images, using an additional axis projection. The implementation using 2D and 3D data will be compared to evaluate pros and cons of adding complexity to the system. Tuples of possible relations will be studied to avoid an explosion of number of relational combinations.

The laser data collected in the experiment will be used to improve the quality of the input data, as it contains the three-dimensional shape of the objects and its GPS position. Point clouds will also be used to create a 3D version of the scene. The delegation framework [Doherty *et al.*, 2013] is planned to be used in the future, in order to get collaboration between agents.

It is expected that the integration of qualitative and quantitative data, as well as robot's and human's abilities increase the robustness of the system, enabling all involved in a mission to describe and understand an environment to reach their goals.

Acknowledgments

This work was carried out with the support of CISB, Swedish-Brazilian Research and Innovation Center, and Saab AB. The author also would like to thank the scholarship provided by CAPES, and the staff of Autonomous Unmanned Aircraft Systems Technologies Lab from Linköping University for sharing the experiments data. Paulo Santos acknowledges support from IBM-FAPESP Proc. 2016/18792-9.

References

[Allen, 1990] James F Allen. Maintaining knowledge about temporal intervals. In *Readings in qualitative reasoning about physical systems*, pages 361–372. Elsevier, 1990.

[Balbiani *et al.*, 2002] Philippe Balbiani, Jean-François Condotta, and Luis Fariñas Del Cerro. Tractability results in the block algebra. *Journal of Logic and Computation*, 12(5):885–909, 2002.

[Burgard *et al.*, 2000] Wolfram Burgard, Mark Moors, Dieter Fox, Reid Simmons, and Sebastian Thrun. Collaborative multi-robot exploration. In *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, volume 1, pages 476–481. IEEE, 2000.

[Chen *et al.*, 2015] Juan Chen, Anthony G Cohn, Dayou Liu, Shengsheng Wang, Jihong Ouyang, and Qiangyuan Yu. A survey of qualitative spatial representations. *The Knowledge Engineering Review*, 30(1):106–136, 2015.

[Doherty *et al.*, 2010] Patrick Doherty, Jonas Kvarnström, Fredrik Heintz, David Landén, and P-M Olsson. Research with collaborative unmanned aircraft systems. In *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2010.

[Doherty *et al.*, 2013] Patrick Doherty, Fredrik Heintz, and Jonas Kvarnström. High-level mission specification and planning for collaborative unmanned aircraft systems using delegation. *Unmanned Systems*, 1(01):75–119, 2013.

[Doherty *et al.*, 2016] Patrick Doherty, Jonas Kvarnström, Piotr Rudol, Marius Wzorek, Gianpaolo Conte, Cyrille Berger, Timo Hinzmann, and Thomas Stastny. A collaborative framework for 3d mapping using unmanned aerial vehicles. In *PRIMA 2016: Principles and Practice of Multi-Agent Systems*, volume 9862, pages 110–130, Cham, 2016. Springer International Publishing.

[Frank, 1996] Andrew U Frank. Qualitative spatial reasoning: Cardinal directions as an example. *International Journal of Geographical Information Science*, 10(3):269–290, 1996.

[Guesgen, 1989] Hans Werner Guesgen. *Spatial reasoning based on Allen's temporal logic*. International Computer Science Institute Berkeley, 1989.

[Murphy, 2016] Robin R Murphy. Emergency informatics: using computing to improve disaster management. *Computer*, 49(5):19–27, 2016.

[Randell *et al.*, 1992] David A Randell, Zhan Cui, and Anthony G Cohn. A spatial logic based on regions and connection. *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR 1992)*, 92:165–176, 1992.

[Renz, 2002] Jochen Renz. *Qualitative spatial reasoning with topological information*. Springer-Verlag, 2002.

[Santos *et al.*, 2015] Paulo E Santos, Gérard Ligozat, and Marjan Safi-Samghabad. An occlusion calculus based on an interval algebra. In *Intelligent Systems (BRACIS), 2015 Brazilian Conference on*, pages 128–133. IEEE, 2015.

[Santos *et al.*, 2016] Paulo E Santos, Murilo F Martins, Valquiria Fenelon, Fabio G Cozman, and Hannah M Dee. Probabilistic self-localisation on a qualitative map based on occlusions. *Journal of Experimental & Theoretical Artificial Intelligence*, 28(5):781–799, 2016.

[Schlieder, 1993] Christoph Schlieder. Representing visible locations for qualitative navigation. *N. Piera Carrete & M. G. Singh (Eds.) Qualitative Reasoning and Decision Technologies*, 1993.

Learning to Build Qualitative Scenario Models From Natural Language

Maxwell Crouse, Clifton McFate, & Kenneth Forbus
Qualitative Reasoning Group, Northwestern University
2133 Sheridan Road, Evanston, IL, 60208, USA
mvcrouse@u.northwestern.edu, {c-mcfate, forbus}@northwestern.edu

Abstract

Natural language descriptions are commonly used to communicate situations where qualitative reasoning is relevant, but automatic construction of scenario models from language remains a challenging problem. This paper describes an approach for learning how to construct scenario models from natural language text, using small amounts of language-only training data. Evidence Expressions (EEs), which bridge between the general-purpose outputs of a semantic parser and the formal constructs of a qualitative domain theory, are learned from training examples. These EEs are then retrieved and combined by analogy to create scenarios for more complex problems. We demonstrate the effectiveness of our technique on a set of 4th and 5th grade science test questions.

1 Introduction

Understanding how to extract qualitative models in the process of natural language understanding is an important problem for learning by reading (e.g. [Kuehne & Forbus, 2004; McFate et al. 2016b]) and for using that knowledge in answering questions. For example, Crouse & Forbus [2016] found that 29% of elementary science test questions they examined required using qualitative reasoning to answer. This suggests that shallow understanding systems (e.g. [Khashabi *et al*, 2016; Clark *et al*, 2016; Khot *et al*, 2017]) will not suffice for human-level performance on such tests. Given the breadth of natural language, learning approaches look like the only option.

Learning in this domain is difficult: There do not exist the large annotated corpora needed for modern machine learning methods. Even if there were, such approaches lead to building domain-specific language systems, which do not gracefully extend to other domains. Since many interesting problems combine multiple domains, generality is also an important constraint. Our approach [Crouse *et al*, 2018] is to use a domain-general semantic parser which is extended by learning abductive patterns that connect the general-purpose semantics to the representations needed for reasoning about various domains. These *Evidence Expressions* (EEs) are then applied to new problems via analogy. Importantly, EEs are compositional, in that multiple EEs learned for simpler texts can be combined (via multiple analogies) to provide interpretations for more complex texts. In our prior work, EEs were constructed by connection graph techniques, using as input unannotated natural language question-answer pairs (the classic Geoquery factoid Q/A domain). This paper extends these ideas to handle more complex, multi-sentence training examples, as needed for building scenario models.

We begin by reviewing the relevant background. Then we describe our technique, including both learning and application of the learned knowledge. We evaluate our approach by answering questions requiring model formulation from a set of 4th and 5th grade science questions.

2 Background

2.1 Qualitative Process Theory

Qualitative Process Theory [Forbus, 1984] formalizes processes as the mechanism underlying continuous change. The direct effects of a process (e.g. liquid flow into a tub) are called *direct influences* (represented as i+ and i-), and their indirect effects are called *indirect influences* (represented as qprop/qprop-), e.g. the level of the water in the tub. *Model fragments* are compositional schemas that define types of entities and relationships in the world. They have *participants* which are related by the model fragment, *constraints* among participants that determine when the model should be considered, and *conditions* of activation. When a model fragment is active, its *consequences* hold. These are frequently influences, though other relationships can be consequences as well. Consider for example the model fragment in Figure 1, which describes a contained liquid. The first line defines its name. Lines 2 and 3 define the participants, entities of types `Container` and `ContainedStuff`. They play the role of `containerOf` and `containedObject` respectively in an instantiated model fragment. Line 4 defines a constraint, that this model should

```
1. (isa SimpleContainer ContainedSubstance)
2. (mfTypeParticipant ContainedSubstance
         ?container Container containerOf)
3. (mfTypeParticipant ContainedSubstance
       ?stuff ContainedStuff objectContained)
4. (mfTypeParticipantConstraint SimpleContainer
       (PhaseOf ?stuff Liquid)
5. (mfTypeConsequence SimpleContainer
      (qprop ((QPQuantityFn Pressure) ?stuff)
             (AmountFn ?stuff)))
```

Figure 1. *A simple qualitative model of a contained liquid*

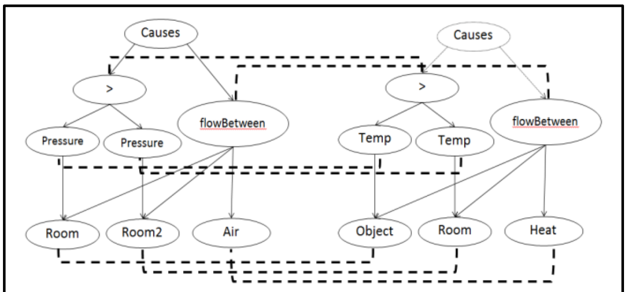


Figure 2. *An example SME mapping*

only be considered for liquids. Finally, Line 5 provides a consequence, that an indirect influence holds between the pressure and amount of contained stuff.

Given a domain theory consisting of a set of model fragments and a scenario description, a *model formulation* algorithm instantiates model fragments based on which preconditions are met by the scenario. In traditional QR, programs that provide structural descriptions do so in the ontology of the domain theory. Here, the challenge is to learn connections between everyday concepts (e.g. tub) and concepts in the domain theory ontology (e.g. container).

2.2 Semantic Parsing

We use the Explanation Agent NLU (EA NLU) semantic parser [Tomai & Forbus, 2009]. EA NLU is a bottom-up rule-based chart parser that uses a feature based grammar. It uses the NULEX lexicon [McFate & Forbus, 2011] and Fillmore et al's [2001] FrameNet. FrameNet ties words to a semantic schema and annotates how semantic roles are bound to arguments in syntactic patterns.

As an example, the word *change* evokes the `Cause_change` frame which has semantic roles that include `Initial_category`, and `Final_category`. When used in the sentence "The snow changes to water.", the first noun phrase is the `Initial_category`, and the prepositional phrase is the `Final_category`. These patterns of role bindings (called valence patterns) are stored in the ontology in templates that get bound in the grammar. In our system, both the grammar and semantic templates are represented in the Cyc ontology [Matuszek *et al*, 2006]. Semantic ambiguities are represented via mutually exclusive choice sets, e.g., the word *ball* in a sentence would lead to including a choice set for a toy versus a dance in the parser's interpretation. Our system operates over these choices to reason about which combination of them would lead to a good qualitative model.

2.3 Analogy

We use the Structure Mapping Engine [Forbus *et al*, 2017], a computational implementation of Gentner's [1983] Structure Mapping Theory. Structure mapping aligns hierarchical structured representations (predicate calculus) according to the principles of SMT, that each element in a case may match with at most one in the other and that the children of matched elements also match. After alignment, structure in one case that is missing from the other can be inferred by analogy (*candidate inferences*).

Consider Figure 2 which shows a mapping between a model of heat flow and a model of air flow. In the base (left), a difference in air pressure between two rooms causes air to flow from one to the other. In the target (right), a difference in temperature exists between a hot brick and cool room. As in the base, there is a flow process between the two entities. When aligned, the ordinal relations and flow process match, allowing the inference that, as in the case of air flow, a quantity difference drives heat flow.

We use the MAC/FAC model of analogical retrieval [Forbus *et al*, 1995] and the SAGE model of analogical generalization [McLure *et al*, 2015]. MAC/FAC is a two-phase model of analogical retrieval that uses a cheap preliminary feature-vector match to generate candidates for its second stage, which uses SME to select the most similar retrieval. SAGE incrementally accumulates examples in a *generalization pool*, a kind of case library, which contains both examples and automatically constructed generalizations. When a new example arrives, it uses MAC/FAC to find the most similar item, and assimilates them if they are sufficiently similar. The assimilation process produces (or updates) a generalization, where the probability of each statement in it depends on the number of examples that contribute corresponding statements. In Figure 2, for instance, the flow-between statement would have probability 1, whereas the temperature/pressure statements would have probability 0.5.

3 Our Technique

Following Crouse *et al* [2018], we represent the mapping between lexical semantics and model fragment ontology as *Evidence Expressions*. An `evidenceForExpression` statement (EE) can be viewed as an abductive rule, where the antecedents are semantic choices and the consequents are a task-specific logical form. In this case, consequents are the participant relations and conditions necessary to instantiate a set of model fragments.

1. (isa ?self NaiveMeltingProcess)
2. (mfTypeParticipant NaiveMeltingProcess ?thing-melting SolidTangibleThing focusOf)
3. (mfTypeParticipant NaiveMeltingProcess ?sub ChemicalCompoundTypeByChemicalSpecies substanceOf)
4. (mfTypeParticipantConstraint NaiveMeltingProcess (substanceOfType ?thing-melting ?sub))
5. (mfTypeParticipantConstraint NaiveMeltingProcess (relationAllInstance freezingPoint ?sub ?m-temp))
6. (mfTypeCondition NaiveMeltingProcess (qGreaterThan (TemperatureFn ?thing-melting) ?m-temp))
7. (mfTypeConsequence NaiveMeltingProcess (qGreaterThan (LiquidGenerationRateFn ?self) 0))
8. (mfTypeConsequence NaiveMeltingProcess (qprop (LiquidGenerationRateFn ?self) (TemperatureFn ?thing-melting)))
9. (mfTypeConsequence NaiveMeltingProcess (i- (AmountOfFn ?sub Solid-StateOfMatter ?thing-melting)(LiquidGenerationRateFn ?self)))
10. (mfTypeConsequence NaiveMeltingProcess (i+ (AmountOfFn ?sub Liquid-StateOfMatter ?thing-melting) (LiquidGenerationRateFn ?self)))

Figure 3. *A qualitative model of melting*

A student is investigating changes in the states of matter. The student fills a graduated cylinder with 50 milliliters of packed snow. The graduated cylinder has a mass of 50 grams when empty and 95 grams when filled with the snow. The packed snow changes to liquid water when the snow is put in a warm room. Which statement best describes this process?

Figure 4. *An example question involving "melting"*

EEs learned during training can be applied to novel questions via analogy. More concretely, question semantics are aligned to the antecedents of an EE retrieved via MAC/FAC, and the consequent of the EE is then instantiated with the entities of the novel question via analogical inference. Missing antecedents are allowed at a cost and a recombination algorithm determines the smallest set of EEs whose antecedents cover the entirety of the question semantics [Crouse *et al*, 2018].

An example will make this clearer. Figure 3 shows a simple model for melting. Informally, the requirements for activation (lines 2-6) are that there is a solid (line 2), and that its temperature is greater than its melting point (line 5 and 6). Now consider the question described in Figure 4. That packed snow is a solid is not explicitly stated, nor is the initial ordinal relationship of its temperature to the melting point of water. That the room is warm is relevant, since that could lead to a heat flow, and thus a melting. That is consistent with the snow changing to liquid. Finally, there are, from the standpoint of model formulation, irrelevancies, e.g. that a student is conducting an investigation, although the savvy student would read "changes in the state of matter" as a hint. We use this example to illustrate how our technique works.

3.1 Training

Our technique takes as inputs a natural-language scenario paired with the active process of the scenario, drawn from science test questions. For the problem in Figure 4, it would be given the text paired with the word "melting" which refers to the `NaiveMeltingProcess` model fragment. This results in an initial set of model fragments consisting of `NaiveMeltingProcess`, any model fragment participants, as well as any fragments that they depend on. This retrieved set will be referred to as our *target set of expressions*.

EA NLU interprets the scenario. Figure 5 shows a partial set of EA choices for the text in Figure 4. In Figure 5, each word has its alternative predicate calculus interpretations listed as sub-bullet points. The first step is to map relevant elements of the natural language scenario to models and their conditions. This proceeds in three phases: The textual semantics and model are aligned in an *initial matching*. Relevant aspects of the semantics are also found through *stability analysis*, and finally *Steiner tree connecting semantics* are found to bridge disjointed antecedents. Figure 6 illustrates this process.

3.1.1 Initial Matching

Let T be the target set of logical expressions (model fragments) and S be the set of semantic choices (e.g. 1a, 2a-c, 3a-b, and 4a in Figure 5). The initial alignment step operates over the complete bipartite graph $G = (S, T, S\times T)$. A conflict-free matching M in G is a set of vertex-disjoint edges (one-to-one correspondence) such that no two pairs of edges can have expressions from T that conflict in C. In other words, if there is a conflict pair $(t_k, t_l) \in C$ then M cannot simultaneously contain a pair of edges like (s_i, t_k) and (s_k, t_l).

This negative-disjointness constraint makes the matching problem NP-Hard [Darmann *et al*, 2011]. For efficiency, we use a local search procedure that starts from a promising

1. "graduated cylinder"
    1a. (isa graduated-cylinder751136 GraduatedCylinder)
2. "warm"
    2a. (ambientTemperature room752597 Warm)
    2b. (temperatureOfObject room752597 Warm)
    2c. (temperatureOfObject room752597 Warm)
        – 2b and 2c are justified by different parse trees
3. "water"
    3a. (isa water751940 (LiquidFn Water))
    3b. (isa water751940 Water)
4. "snow"
    4a. (isa snow751995 SnowMob)

Figure 5. *A subset of the semantics produced for our example question. 2b and 2c shows a situation where the same semantics are justified by different parse trees.*

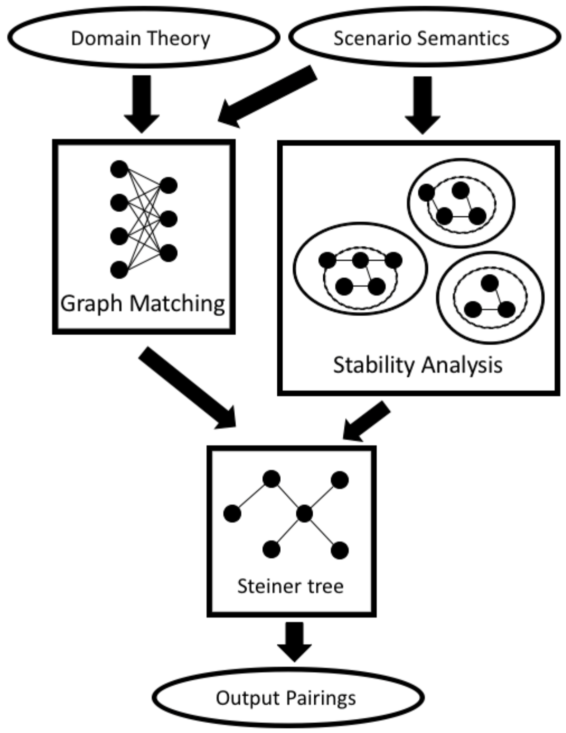


Figure 6. *Training Diagram*

candidate solution and moves amongst better neighboring solutions until it can no longer improve. For a conflict-free matching M in G, it expands outwards to all conflict-free matchings that differ by at most two edges from M.

The score of a conflict-free matching M is determined by three properties: ontological alignment, loose structural overlap, and conflict avoidance. We describe each property and how they are combined next.

Ontological Alignment

The Cyc ontology comes with a number of higher-order relations that relate concepts at an abstract level. For example, functionCorrespondingPredicate indicates that a given function and predicate amount to the same relationship, holding for pairs like temperatureOfObject and TemperatureFn. We use a set of such relationships to estimate expression similarity.

Loose Structural Overlap

We employ a simple, loose measure of structural similarity inspired by the concept of similarity flooding [Melnik *et al*, 2002], due to the distance between the concepts involved.

We first define both sets of logical expressions, S and T, as hypergraphs. Let $H(T)$ be the hypergraph of T where hyperedges are expressions and vertices are the entities and variables contained in those expressions. For instance, the expression on Line 10 of Figure 3, would be a hyperedge connecting ?sub, ?thing-melting, and ?self. By considering these sets to be hypergraphs, we can define a notion of neighbors in these hypergraphs.

For a particular expression t_i, we consider the neighboring expressions to be all those expressions with a distance from t_i that is less than a given constant k. In our experiments, we use $k = 3$. We define the distance between two expressions t_i and t_j to be the length of the shortest path in $H(T)$ connecting *any* two entities of t_i and t_j. The loose-structural overlap score of expressions s_i and t_j is 1 if both expressions have neighbors in $H(G)$ that are also in M and 0 otherwise.

Conflict Avoidance

Our system gives preference to edges that lead to the fewest conflicts overall. Conflicts in the semantic interpretation arise from choices to resolve ambiguities that would lead to logical inconsistencies (i.e. a river bank is not a place to deposit money). The normalized conflict-avoidance score $c(s_i,t_j)$ is the number of edges in G that do *not* conflict with edge (s_i,t_j) divided by the cardinality of G

Calculating and Using the Score

Let $o(s_i,t_j)$ be the ontological alignment score, $s(s_i,t_j)$ be the loose structural overlap score, and $c(s_i,t_j)$ be the conflict avoidance score. Then, the overal score of a matching M between S and T is given as:

$$\sum_{(s_i,t_j)\in M} \alpha * o(s_i,t_j) + \beta * s(s_i,t_j) + \gamma * c(s_i,t_j)$$

where α, β, and γ are the preferences for each property of the matching. In our experiments, we have $\alpha = 0.4, \beta = 0.4, \gamma = 0.2$ which were values chosen simply because conflict-avoidance was found to be important, but less important than the other two features.

The local-search algorithm obtains a first-pass candidate solution by greedily finding a matching M that maximizes for only ontological alignment. From that initial matching M, the algorithm begins to iteratively improve its solution by exploring the conflict-free matchings that differ by at most two edges from M until it can no longer find a matching with a higher score as determined by the equation above. When a new matching with a higher score is found, the algorithm repeats from that matching, otherwise it returns M.

The final matching contains pairings of expressions in S with expressions in T. Some of those pairs include type constraints. For example, Line 3b of Figure 5 ("water" as Water) might match with Line 3 of Figure 3 (the participant with ChemicalCompoundTypeByChemicalSpecies. From those type constraint pairings, our approach extracts variable bindings (e.g. ?sub with water751940) and instantiates the T with those bindings. The result is a subset of the instantiated T that includes only participants, participant constraints, and conditions, that is conditioned on the expressions in S that were used in M.

```
(temperatureOfObject room752597 Warm)      - from matching
(isa water751940 Water)                    - from matching
(isa snow751995 SnowMob)                   - from matching
(isa change2037026 StateChangeEvent)       - from stability analysis
(isa put2038224 PuttingIntoAState)         - from stability analysis
(isa room752597 RoomInAConstruction)       - from stability analysis
```

Figure 7. *Antecedents after matching and stability analysis*

3.1.2 Stability Analysis

Often there are salient features of scenario types that are not indicated by an ontological match. For instance, situations regarding gravity often involve something falling.

Our system creates a SAGE generalization pool for each process type (melting, freezing, gravity, etc.). The generalization pools are populated by the complete sets of semantics for all training instances of that type (e.g. each question about melting is parsed, and its semantics put into a case which is then added to a generalization pool for melting). After the matching is complete, the generalization pool for the given question type is retrieved and is used to assign probabilities to all of the expressions in *S*. The top 3 most probable expressions from *S* that do not conflict with already selected expressions from the matching step are added to our antecedents. Figure 7 shows the antecedents (each of which are predicate calculus interpretations of phrases from our training paragraph) after both the matching and stability analysis and indicates the source of each antecedent.

3.1.3 Steiner Tree Connecting Semantics

At this point, antecedents have been drawn from both the initial matching and generalized stable structures. Referring back to Figure 7, it is clear that the semantics selected thus far are disconnected from one another. We would like to incorporate the context of the question that includes and connects all of those expressions. We pose this as the problem of finding a conflict-free Steiner tree through the hypergraph *H*(*S*) that connects all the entities seen in our selected set of choices from *S* (in Figure 7, those entities would be snow751995, water751940, room752597, etc).

The minimum Steiner tree problem in graphs is the problem of finding the minimum cost tree in a graph *G* that connects a given subset of its vertices. While there are approximation algorithms for the Steiner tree problem e.g. [Agrawal *et al*, 1995], there are no approximation algorithms that take into account negative-disjointness constraints. We use a simple extension to a 2-approximation algorithm for the minimum Steiner tree problem (though we make no optimality guarantees) to ensure it produces a non-conflicting set of semantic choices. While our algorithm is not guaranteed to result in the antecedents becoming fully connected, it appears to work well in practice. It also allows for the straightforward addition of coreference resolution, by extending the set of semantics to include coreference expressions taking two coreferable entities as arguments. The Steiner tree algorithm can connect entities across sentences when necessary by using those expressions.

```
(evidenceForExpression
 (and (isa snow751995 SolidTangibleThing)
      (isa water751940 ChemicalCompoundTypeByChemicalSpecies)
      (substanceOfType snow751995 water751940)
      (relationAllInstance freezingPoint water751940 abduced-temp12)
      (qGreaterThan (TemperatureFn snow751995) abduced-temp12)
 (and (temperatureOfObject room752597 Warm)      - from matching
      (isa water751940 Water)                    - from matching
      (isa snow751995 SnowMob)                   - from matching
      (isa change2037026 StateChangeEvent)       - from stability analysis
      (isa put2038224 PuttingIntoAState)         - from stability analysis
      (isa room752597 RoomInAConstruction)       - from stability analysis
      ((VerbRelFn be) water2037198 snow751995) - from Steiner tree
      (fe_effect put2038224 room752597)          - from Steiner tree
      (fe_Cause put2038224 snow751995)           - from Steiner tree
      ((IBPFn parts) room752597 change2037026) - from Steiner tree
      (objectActedOn change2037026 snow751995))) - from Steiner tree
```

Figure 8. *The final EE (sources of antecedents are to the right)*

3.1.4 Storing Cases

The pairing of choices from *S* and model conditions will be stored in a case library as an evidenceForExpression statement as its own case. Figure 8 shows the final EE produced for the training question in Figure 4, as well as sources for each of the antecedents of the EE. The consequent of the EE are the forms required to instantiate the model fragment (i.e. its participants and constraints).

3.2 Testing

Following training, we have a set of evidenceForExpression statements whose antecedents are question semantics and whose consequents are activation requirements of model fragments (e.g. participants of the correct type and relations between them).

3.2.1 Retrieval and Instantiation

Given a new scenario, it is first interpreted by EA NLU. The complete set of undisambiguated semantics forms a case. MAC/FAC then retrieves the five most similar EEs (i.e. the ones with the most antecedent overlap) from the case library of EEs. The consequents of an EE are inferred by analogical inference and bound with the variables from the scenario interpretation.

EEs are abducible in that only a subset of the antecedents are required to infer the consequent activation conditions. An initial ranking of the EEs is given by the number of abduced antecedents. To prevent over-eager application of EEs to a given scenario, our technique only considers those EEs with at least as many antecedents satisified as abduced.

3.2.2 Composing EEs

The ordered EEs are input into a slight variant of the query composition algorithm of Crouse *et al* [2018]. The modification excludes conflict counts from being considered in the score of an EE, which turns the composition algorithm into a greedier coverage-focused algorithm (i.e. an algorithm that looks for EEs covering as many semantic choices as possible, without regard to how many semantic choices the selected EEs would rule out through choice-set constraints). We briefly recap the algorithm here.

The composition algorithm takes a set of instantiated EEs and a set of semantics *S* to be covered. The antecedents of each EE are elements of *S*, where *S* is the set of semantic choices for the current scenario. The algorithm iteratively selects the EE whose antecedents cover as much of *S* is possible, removing semantic choices from *S* that conflict with selected EEs as it goes. When *S* is empty (because the elements of *S* were either covered by some selected EE or conflicted with the antecedents of some selected EE) the algorithm returns the consequents of every selected EE. The appeal of the algorithm is that it treats the EE selection process as a coverage problem, producing the smallest set of consequents that reflects as much of the semantics as is possible.

3.2.3 Model Formulation and Question-Answering

The output of the composition algorithm is a set of activation conditions that can be used to instantiate model fragments relevant to the scenario at hand. Our technique uses this model formulation algorithm to instantiate all applicable model fragments and collects the resultant facts about the scenario into a set of model facts *F*.

Our method is evaluated on two types of scenario questions: standard questions (e.g. those that end in a question mark like "What process occurred?") and fill-in-the-blank questions (e.g. "When ice melts it ____"). To handle standard questions, we first filter out all expressions from *F* that do not have a positive alignment score (like in training) with our question semantics. Then, for each answer option, the semantics for the answer are matched against *F* using the same matching procedure described in training. For fill-in-the-blank questions the process is largely the same. For each answer option, the semantics of the question *and* answer are matched against *F* using the matching procedure from training, and the highest scoring answer of the question-answer pairs is output as the answer to the question. If no models can be instantiated, then no answer will be chosen.

4 Experimental Evaluation and Discussion

We evaluate our approach on a set of 45 questions from 4th and 5th grade elementary science tests. This set of questions was extracted by a script that searched across a large set of science test questions (collected by the Allen Institute for Artificial Intelligence). The questions the script returned had keywords associated with the model fragments outlined in Crouse and Forbus' [2016] science test analysis. Furthermore, the questions were restricted to those involving reasoning about a scenario, not questions involving definitions or taxonomies. Future work will involve determining all of the phenomena in the elementary science tests that can be representable by QR models.

Question Type	Correct / Total	Percent
All	26 / 45	58%
Model Formulation	19 / 26	73%
Model Reasoning	7 / 19	37%

Table 1. *Average Performance*

The questions our system was designed to handle were those only requiring model formulation. Those involving more complex QR techniques like qualitative simulation or differential qualitative analysis were out of the scope of this work. We categorized those questions requiring only model formulation as *model formulation* questions and those requiring additional reasoning on top of relevant instantiated model fragments generated by model formulation (i.e. all others) as *model reasoning* questions.

We evaluated using 5-fold cross validation (i.e. 36 questions for training, 9 questions for testing per fold). Table 1 shows our overall average performance on all questions, model formulation questions, and model reasoning questions. Random guessing on this dataset would lead to 25% correct.

The performance gap between model formulation and model reasoning questions is quite substantial. This comes as no surprise, given that our approach is not yet learning how to answer more reasoning-intensive questions during training. This gives one immediate avenue for future work, which is to extend our approach to learn the reasoning needed to answer the more advanced questions through only question-answer pairs.

4 out of 7 errors for model formulation questions were due to phrasings outside of the handling of our approach. For example, the question, "Which type of force requires contact between two objects for one to …" requires one to know that "two objects" implies there are two distinct objects in the scenario. Our approach only identifies one object from that scenario, and thus cannot successfully instantiate a model fragment for friction. The remaining 3 errors were due to inadequate training data, where our approach hadn't seen a scenario similar enough to formulate the correct model to answer the question.

The model reasoning questions gave our approach much more difficulty. Apart from requiring more sophisticated reasoning techniques, the language of the questions tended to be more complex. Accordingly, for 4 out of 19 questions our approach found the correct model fragment to use, but instantiated it incorrectly. For 6 out of 19 questions the wrong model fragment was selected, while in another 6 out of 19 questions the correct model fragment was instantiated. The unfortunate last source of issues were processes that were never seen during training. Our corpus only had two questions revolving around fluid displacement, and both of those questions were found in the same fold.

5 Related Work

Barbella & Forbus [2011] introduced *analogical dialogue acts* (ADAs), which formalize the roles played by individual utterances in instructional analogies. Their approach used the ADAs recognized from the semantic parse of an instructional text to build structured cases that were then compared with SME. Their system used inferences from these analogies to interpret and answer questions. Our approach also uses analogical inferences to construct an interpretation of text (scenario model), however our system goes a step further in that EEs are learned from natural language while ADAs were recognized with manually constructed rules. They also used dynamic case construction [Mostek *et al*, 2000] to automatically extended their cases with pertinent background facts from the knowledge base, which may be a useful technique to incorporate into our system to validate activation conditions which may be available as stored knowledge.

Barbella & Forbus [2015] further present a method for constructing coherent cases from text which is similar to our goal of extracting relavent semantics. One way our approaches differ is that our relevence condition is overlap with a pre-existing model while theirs finds facts related to a seed from the text. Their approach is complementary to ours and could be used to segment a corpus into cases from which our approach could build interpretations.

Chang [2016] combined natural language understanding, spatial reasoning, and analogical reasoning to interpret instructional analogies. These analogies could be used to learn qualitative knowledge. Of particular relevance to our work was the use of visual representations to disambiguate natural language. Their work used the CogSketch sketch understanding system [Forbus *et al*, 2011] to represent sketches with the Cyc ontology. EA NLU semantic choice-sets were disambiguated by selecting those choices that were most related to the outputs of the sketch understanding system. This could be incorporated naturally into our work as part of the ontological features our approach uses during the matching step.

Khashabi *et al* [2017] introduced the notion of *essential question terms,* which were terms absolutely critical to the understanding of a particular question. They showed that without those terms, human performance on science test questions dropped significantly. This is related to our approach which learns the essential components of a scenario needed to infer the activation conditions of a particular model fragment (EEs).

Khot *et al* [2017] introduced a method for answering complex, compositional science test questions from OpenIE extracted knowledge bases. They posed the problem of multiple-choice question-answering as a search for an optimal subgraph connecting a question and answer through the knowledge base. The types of questions this system was designed for were compositional factoid questions, which likely makes their system complementary to ours.

Fan and Porter [2004] introduced Loose-speak, an interpreter intended to fix misalignments commonly seen between the queries of novice users of a knowledge base and the knowledge base being queried. It was equipped with a variety of features for determining when an expression in the knowledge base was likely the intended expression of the user, some of which are similar to the ontological alignment features of our work.

6 Conclusions and Future Work

We have described an implemented system that adapts a domain-general semantic parser to build qualitative scenario models. Our system uses these models to answer elementary science test questions. Our approach builds on prior work by Crouse *et al* [2018] and uses their EE formalism for cases that are retrieved and applied by analogy to adapt the parser's outputs. The primary contribution of this work is a novel technique which associates the relevant semantics of a scenario to qualitative model fragments and applies these associations by analogy to construct a scenario model. This approach goes significantly beyond prior work in its ability to build models for multi-sentence scenario questions and in providing a richer framework for judging ontological similarity.

There are three clear directions for future work. First, our approach only utilized a small number of qualitative models and thus it was limited in scope. Future work will use the mapping between the core elements of QP theory and FrameNet produced in [McFate & Forbus, 2016a] to extract qualitative models *a la* [McFate & Forbus, 2016b] and broaden the range of model fragments our system can utilize. Second, we will extend the reasoning capabilities of our system to handle a larger range of questions. Finally, we do not currently generate natural language responses, nor natural language explanations for the answers given. Generating both would set the stage for teaching and correcting the system via interactive dialogue.

Acknowledgements

This research was supported by the Intelligent and Autonomous Systems Program of the Office of Naval Research

References

[Agrawal *et al*, 1995] Agrawal, A., Klein, P., & Ravi, R. (1995). When trees collide: An approximation algorithm for the generalized Steiner problem on networks. *SIAM Journal on Computing*, *24*(3), 440-456.

[Barbella & Forbus, 2011] Barbella, D. and Forbus, K. (2011). Analogical Dialogue Acts: Supporting Learning by Reading Analogies in Instructional Texts. *In Proc. AAAI-11*, San Francisco, CA.

[Barbella & Forbus, 2015] Barbella, D. and Forbus, K. (2015). Exploiting Connectivity for Case Construction in Learning by Reading. *In Proc. of the Third Annual Conference on Advances in Cognitive Systems.*

[Chang, 2016] Chang, M. (2016). Capturing Qualitative Science Knowledge with Multimodal Instructional Analogies (Doctoral dissertation, Northwestern University).

[Clark *et al*, 2016] Clark, P., Etzioni, O., Khot, T., Sabharwal, A., Tafjord, O., Turney, P. D., & Khashabi, D. (2016, February). Combining Retrieval, Statistics, and Inference to Answer Elementary Science Questions. *In Proc. AAAI-16* (pp. 2580-2586).

[Crouse & Forbus, 2016] Crouse, M., & Forbus, K. D. (2016). Elementary School Science as a Cognitive System Domain: How Much Qualitative Reasoning is Required?. *Advances in Cognitive Systems*, *19*.

[Crouse *et al*, 2018] Crouse, M., McFate, C., & Forbus, K. (2018). Learning from Unannotated QA Pairs to Analogically Disambiguate and Answer Questions. In *Proc. AAAI-18*

[Darmann *et al*, 2011] Darmann, A., Pferschy, U., Schauer, J., & Woeginger, G. J. (2011). Paths, trees and matchings under disjunctive constraints. *Discrete Applied Mathematics*, *159*(16), 1726-1735.

[Fan & Porter, 2014] Fan, J., & Porter, B. (2004, July). Interpreting loosely encoded questions. In *Proc AAAI-14* (pp. 399-405).

[Fillmore *et al*, 2001] Fillmore, C. J., Wooters, C., & Baker, C. F. (2001). Building a large lexical databank which provides deep semantics. In *PACLIC-15* (pp. 3-26).

[Forbus, 1984] Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence*, 24, 85-168.

[Forbus *et al*, 1995] Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive science*, *19*(2), 141-205.

[Forbus *et al*, 2011] Forbus, K., Usher, J., Lovett, A., Lockwood, K., & Wetzel, J. (2011). CogSketch: Sketch understanding for cognitive science research and for education. *Topics in Cognitive Science*, *3*(4), 648-666.

[Forbus *et al*, 2017] Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2017). Extending SME to Handle Large-Scale Cognitive Modeling. *Cognitive Science*, *41*(5), 1152-1201.

[Gentner, 1983] Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science*, *7*(2), 155-170.

[Khashabi *et al*, 2016] Khashabi, D., Khot, T., Sabharwal, A., Clark, P., Etzioni, O., & Roth, D. (2016, July). Question answering via integer programming over semi-structured knowledge. *In Proc. IJCAI*-16. AAAI Press.

[Khashabi *et al*, 2017] Khashabi, D., Khot, T., Sabharwal, A., & Roth, D. (2017). Learning What is Essential in Questions. In *Proc. CoNLL 2017)* (pp. 80-89).

[Khot *et al*, 2017] Khot, T., Sabharwal, A., & Clark, P. (2017). Answering Complex Questions Using Open Information Extraction. In *Proc. ACL-17*.

[Matuszek *et al*, 2006] Matuszek, C., Cabral, J., Witbrock, M. J., & DeOliveira, J. (2006, March). An Introduction to the Syntax and Content of Cyc. In *AAAI Spring Symposium* (pp. 44-49).

[Mostek *et al*, 2000] Mostek, T., Forbus, K, and Meverden, C. (2000). Dynamic case creation and expansion for analogical reasoning. *In Proc. of AAAI-2000*. Austin, TX.

[McFate & Forbus, 2011] McFate, C. J., & Forbus, K. D. (2011, June). NULEX: an open-license broad coverage lexicon. In *Proc. NACL-HLT-11*

[McFate & Forbus, 2016a] McFate, C., & Forbus, K. (2016). *An Analysis of Frame Semantics of Continuous Processes*. *In Proc. CogSci-16*

[McFate & Forbus, 2016b] McFate, C., & Forbus, K. (2016). Scaling up Linguistic Processing of Qualitative Processes. *In Proc. Advances in Cognitive Systems 2016.*

[McLure *et al*, 2015] McLure, M. D., Friedman, S. E., & Forbus, K. D. (2015, January). Extending Analogical Generalization with Near-Misses. In *Proc. AAAI-15* (pp. 565-571).

[Melnik *et al*, 2002] Melnik, S., Garcia-Molina, H., & Rahm, E. (2002). Similarity flooding: A versatile graph matching algorithm and its application to schema matching. In *Data Engineering, 2002. Proceedings. 18th International Conference on* (pp. 117-128). IEEE.

[Schoenick *et al*, 2017] Schoenick, C., Clark, P., Tafjord, O., Turney, P., & Etzioni, O. (2017). Moving beyond the turing test with the allen ai science challenge. *Communications of the ACM*, *60*(9), 60-64.

[Tomai & Forbus, 2009] Tomai, E., & Forbus, K. D. (2009, March). EA NLU: Practical Language Understanding for Cognitive Modeling. In *FLAIRS Conference 2009*.

A Qualitative Color Harmony Theory and its Application to Art Images

Lledó Museros[a], Ismael Sanz[a], Zoe Falomir[b], and Luis Gonzalez-Abril[c]
[a]Universitat Jaume I, Castellón, Spain
{museros,isanz@uji.es}
[b]Universität Bremen, Bremen, Germany
zfalomir@uni-bremen.de
[c]Universidad de Sevilla, Sevilla, Spain
luisgon@use.es

Abstract

This paper presents a qualitative color harmony (QCharm) theory based on a Qualitative Color Descriptor (QCD) together with the definition of 5 basic color combination operations. The initial QCD is a color naming model validated and customizable by users. Then QCharm using this QCD is able to construct color palettes (or color schemes) of compatible colors. Palettes of 3 colors are constructed and their harmonious quality is predicted using a model defined using data from large datasets such as ColorLovers[1]. Moreover, the paper also presents QArt-Learn, a qualitative model for categorizing paintings in art styles.

1 Introduction

Humans attach labels to relevant features of stimuli/things, so that a classification can be performed and further abstract work can be carried out. Color is one of these essential features. Any technology that aims to reproduce some level of human intelligence, or involves human-machine communication regarding the environment (for instance, assistive technology or ambient intelligence systems), must be able to successfully recognize and categorize colors.

From the point of view of color naming research, Kay and Regier [2006] found that: (i) color categories appear to be organized around universal color foci, but (ii) naming differences across languages do cause differences in color cognition because color categories are determined at their boundaries by language. Therefore, a cognitive color-naming model should be parameterized not only in a universal way, so that an intelligent agent (robot or ambient intelligent system) can communicate prototypical colors to a human user, but also in a specific way, so that the system adapts to a specific type of user or community which understand colors differently.

A great effort has been done to mimic the *color naming* capabilities of humans. In the literature, color-naming models have been defined using different color spaces. Menegaz et al. [2007] presented a model for computational color categorization and naming based on the CIE Lab color space and fuzzy partitioning. Mojsilovic [2005] presented a computational model for color categorization and naming and extraction of color composition based on the CIE Lab and HSL color spaces. Seaborn et al. [2005] defined fuzzy color categories based on the Musell color space (L*C*H). Liu et al. [2004] converted the dominant color of a region (in HSV space) into a set of 35 semantic color names. Stanchev et al. [2003] defined 12 fundamental colors based on the CIELUV color space and used Johannes Itten's theory of color to define the contrasts light-dark, warm-cold, etc. Corridoni et al. [1998] presented a model for color-naming based on the Hue Saturation and Lightness (HSL) color space and also introduced some semantic connotations, such as warm/cold or light/dark colors. Lammens [1994] presented a computational model for color perception and color-naming based on the CIE XYZ, CIE Lab and NPP color spaces.

Here, the Qualitative Color Description (QCD) [Falomir et al. 2013] model, based on the HSL color space is used. In contrast to the work by Corridoni et al. [1998] and Mojsilovic [2005], the QCD model is kept as simple as possible, since Conway [1992] demonstrated that human beings can only identify a reduced number of color names. Furthermore, the QCD model may be understandable by non-experts in color.

Regarding *color customization*, [Soto-Hidalgo et al. 2010] proposes a model for customization in fuzzy color spaces. However, to the best of our knowledge our model was the first customization based on a qualitative color model [Sanz et al. 2015]. The model has been validated with an experiment designed to determine if the color descriptions it produces are close to those of human beings. Also, [Sanz et al. 2015] demostrates that color is a subjective feature; therefore it is very important to customize the color intervals to a user profile for establishing successful communication. For example, the color name included in a user interface, which can be both written and read aloud by a speech synthesizer application to help users to communicate (even blind and/or deaf users), will be understood by the system and the user as the same colorimetric feature. That is, there will not be ambiguities between the subjective perspective of the user and the color reference system used by the machine.

[1] http://www.colorlovers.com/

On the other hand, *color harmony* has relevant significance in art and design. Color can be a powerful design element if used effectively: with colors you can set a mood, attract attention, energize, or cool down. By selecting the right color scheme, you can create a chic, romantic, modern or elegant ambiance.

Color harmony has been defined [Moretti et al. 2013] as "an intuitively simple idea that can be expressed as a set of colors that look good when seen together". When people are asked about *how much they liked a color combination*, they are asked about their color preference. When people are asked about *how well two colors went together* – regardless the preference –, they are asked about the harmony of colors [Palmer and Griscom 2013]. In this paper, color preference refers to preference for harmonious colors. That is, we are not focusing on individual differences on color preference, but our target is the general public. As it happens in music, not everybody prefers harmonious music --someone would prefer strident music--, but when organizing an event for general public, usually harmonic music is fine for most of the audience.

In the literature, harmonic colors have been defined as sets of colors that hold some special internal relationship that provides a pleasant visual perception. Harmony among colors is not determined by specific colors, but rather by their relative position in a color space. Although a definition of color harmony is challenging, there are many heuristics that have been found to be helpful in choosing sets of harmonious colors. The most widely used are those based on color wheels and templates [Itten 1970], including those for selecting complementary color and split-complementary schemes among others.

Itten [1970] introduced a new kind of color wheel for which color harmony was defined, with an emphasis on hue. Itten's color harmony theory is based on the relative positions of the hues on the color wheel. For example, from the three primary colors of cyan, magenta, and yellow, Itten designed a hue wheel of twelve colors. Itten [1970] claimed that any combination of colors is harmonious if the colors produce neutral gray when mixed together. He referred to complementary colors as a two-color harmony, three-color harmony of hues that form an equilateral triangle, the four-color harmony of hues forming a square, the six-color harmony of a hexagon, etc. His schemes have been widely adopted by artists and designers [Kuang et al. 2014]. Other theories as Munsell's [Munsell 1921] stated that colors are harmonious when they have certain relations in color space, such as, being constant in hue and saturation but varying in lightness. Another common harmonious color approach is to produce a monochromatic color scheme comprising various saturations and lightness levels of a single hue [O'Connor 2014]. The color harmonization operators presented in this paper are also based on these schemes.

The main aim of developing a color harmony theory is to judge how well a set of colors match when viewed as a whole. Therefore, it is useful to develop a way to rate the goodness or harmony of a color scheme or palette, to obtain the degree of pleasantness that harmonious color combinations produce on people. In recent decades, psychologists have studied the user preference about a color combination [Palmer and Griscom 2013, Schloss and Palmer 2011, Ou and Ronnier Lou 2006, Ou et al. 2012].

Moreover, software applications including preferences about color combinations have also appeared. In [Jahanian et al. 2013], Jahanian et al. presents a recommendation system for automatic design magazine covers by generating several alternative designs that can be rated by the user. Nishiyame et al. [2011] present an aesthetic quality classification method for photographs based on a large photo collection with user-provided aesthetic quality scores. All these studies are done based in a context, but our model [Museros et al. 2016] follows the trend presented in [O'Donnovan et al. 2011], an a rating methodology of color palettes is defined in a more generic way, where palettes are rated without a context.

Finally, the QCHarm model developed has been applied to the problem of *art paintings categorization* [Falomir et al. 2018]. Previous research have dealt with the problem of art paintings categorization. Shen [2009] classified paintings by artist using a radial basis function (RBF) neural network classifier. Gatys et al. [2015] presented an artificial neural system that achieved a separation of image content from style using deep neural networks. Karayev et al. [2013] applied deep neural networks trained on object recognition for style recognition in order to classify artworks according to their period. Shamir et al. [2010]; Shamir and Tarakhovsky [2012] automated the recognition of nine painters and three schools of art based on their signature styles. The features they used were those by Orlov et al. [2007] for analyzing microscopy images and the learning techniques applied were weighted nearest neighbor (WNN), and SVMs with linear, polynomial and RBF kernels. Siddiquie et al. [2009] used Boost-based SVMs, an alternate method to select training data instances using AdaBoost for each of the SVM base kernels, to classify painting styles. Li and Chen [2009] presented a measure to assess the aesthetic visual quality of paintings. Mensink and van Gemert [2014] classified paintings in Rijksmuseum according to their author using Fisher vectors of local SIFT descriptors and SVMs. Moreover, wavelet of brush strokes have also been analyzed in paintings to find out artist identity which had very successful results in the case of identifying Vicent van Gogh's paintings [Li et al., 2012].

However, none of the features used in the previous classification works were qualitative concepts. Qualitative descriptors have been proved to be successful in managing incomplete, imprecise and ambiguous information [Cohn and Renz, 2007; Freksa, 2013]. And, moreover, they use linguistic concepts which align with human perception and can be easily used to generate explanations and give feedback to the users. Few research works, such as that by Yelizaveta et al. [2005] used semantic categories (i.e. warm, cold) and color names proving to be effective for painting retrieval in databases. The QArt-Learn approach is defined for categorizing paintings into art styles implemented using machine learning techniques such as k-Nearest Neighbour (k-NN) and Support Vector Machines (SVMs). It combines qualitative color palettes with average color features and analyses their performance. For testing the QArt-Learn approach, the Painting-91 dataset by Khan et al. [2014] actively applied in the literature was used.

2 Overview of the QCD

The QCD model [Falomir et al. 2013] defines a reference system in the HSL color space for qualitative color description, which is built according to Figure 1 and defined as:

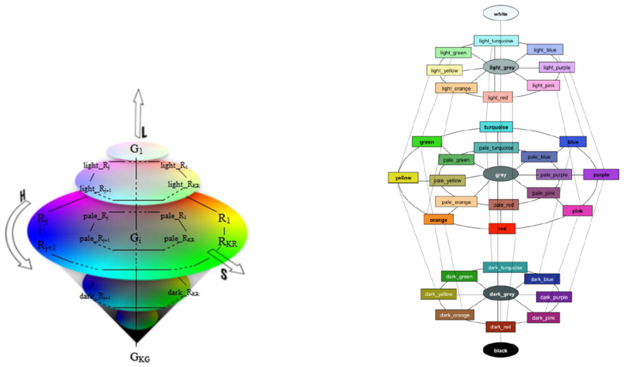
$$QC_{RS} = \{uH, uS, uL, QC_{NAME1..5}, QC_{INT1..5}\}$$

where uH is the unit of Hue; uS is the unit of Saturation; uL is the unit of Lightness; $QC_{NAME1..5}$ refers to the color names; and $QC_{INT1..5}$ refers to the intervals of HSL coordinates associated with each color. The chosen QC_{NAME} and QC_{INT} are:

QC_{NAME1} = {*black, dark grey, grey, light grey, white*}
$QC_{INT1} = \{[0, 20), [20, 30), [30, 50), [50, 75), [75, 100) \in uL \mid \forall\ uH \wedge uS \in [0, 20]\}$
QC_{NAME2} = {*red, orange, yellow, green, turquoise, blue, purple, pink*}
$QC_{INT2} = \{(335, 360] \wedge [0, 20], (20, 50], (50, 80], (80, 160], (160, 200], (200, 260], (260, 300], (300, 335] \in uH \mid uS \in (50, 100] \wedge uL \in (40, 55]\}$
QC_{NAME3} = {pale-red , pale-orange, pale-yellow, ..., pale-blue, pale-purple, pale-pink}
$QC_{INT3} = \{\forall\ QC_{INT2} \mid uS \in (20, 50] \wedge uL \in (40, 55]\}$
QC_{NAME4} = {light-red, light-orange , light yellow , ..., light blue , light purple , light pink}
$QC_{INT4} = \{\forall\ QC_{INT2} \mid uS \in (50, 100] \wedge uL \in (55, 100]\}$
QC_{NAME5} = {dark red , dark orange , dark yellow , ..., dark blue , dark purple , dark pink}
$QC_{INT5} = \{\forall\ QC_{INT2} \mid uS \in (50, 100] \wedge uL \in (20, 40]\}$

As a baseline, the QCRS was calibrated according to the vision system used.

The QCD model has a relational structure and it can be organized in a conceptual neighborhood diagram (CND) [Freksa 2013] according to how a color can be transformed into another by changing its luminosity, saturation or hue. For example, the colors *red* and *orange* are conceptual neighbors since a continuous change in hue causes a direct transition from *red* to *orange*. However, *blue* and *red* are not conceptual neighbors, since a continuous transformation of hue from blue to red finds other colors in between. A CND for the computational QCD is built and shown in Figure 1(b).



(a) (b)

Figure 1. Diagram for describing QCD: (a) Discretization of the HSL color space; (b) CND corresponding to QCD, w_i correspond to the weights used to calculate the similarity between colors

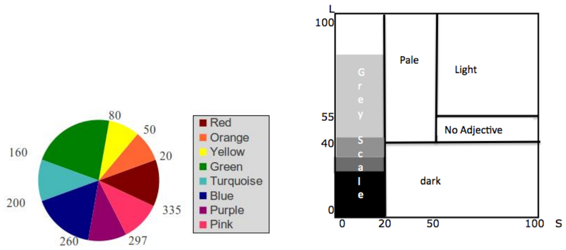
2.1 Validating and Customizing the QCD model

In this section two experiments are presented: the first experiment is intended to validate the QCD model as generating names close to human color perception; and the second one is intended to demonstrate the necessity of customizing the color description based on a user profile.

In order to parameterize the QCD model using a taxonomy of colors as general as possible, an experiment was done in [Sanz et al. 2015] where people were asked to freely determine a name and an adjective (if necessary) for describing a displayed color. From their responses it was observed that the most used names were the colors {*red, orange, yellow, green, turquoise, blue, purple, pink*} and the adjectives {*pale, light, dark*}. With respect to the grey scale, the most used color names were {*black, dark_grey, grey, light_grey, white*}.

Therefore, the QCD was parameterized using a QCRS including 11 basic colors (*black, grey, white, red, orange, yellow, green, turquoise, blue, purple, pink*) and the semantic descriptors *pale*, *light* and *dark*. That is, a total of $5 + 8 \times 4 = 37$ color names were obtained. The thresholds in HSL to determine these prototypical 37 colors were calculated using the AMEVA discretization algorithm [Gonzalez-Abril et al. 2009].

Then, the resulting intervals were used in a test in order to validate this specific parameterization of the QCD model, that is, to test that these intervals divide the HSL color space into the same color names most people use and recognize. To validate this parameterization of the QCD model, a web test similar to the previous one were carried out, but participants could not freely chose a color name, they can only choose among the 37 color names of the model. A total of 545 responses were obtained, and they were used to modify the intervals for color names. As a result of this experiment, the initial intervals (Figure 2) were modified by the intervals shown in Figure 3.



(a) (b)

Figure 2. Diagram describing the QCD parameterization into 5 color sets and 37 color names.

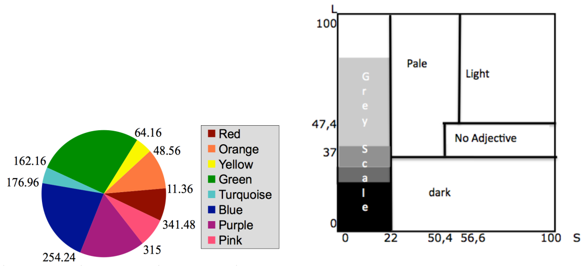


Figure 3. Graphical interpretation of the new color intervals.

To test the assumptions about the subjectivity of color naming and the need for customizability, a second experiment has been carried out including special objects, such as clothes, which can have different colors and shadows. Given an image, the application determines the predominant color of the picture from the set of colors of the QCD. Then it asks the user if she agrees with the name given to the color. If the user agrees, the color name is stored into the database. If the user disagrees, the application asks the user to determine the color name by: (i) freely giving a name to the color, and (ii) selecting one of the color names of the QCD. In order to carry out the tests a database of 37 images of different fabrics was created. For each image its histogram is calculated and the most representative color is the one selected as the color of the fabric. 5 users did the test. The most relevant result is that there is a great variability, which again shows that color is a very subjective perception. Therefore, it is very important to customize a color system able to communicate in an efficient way with users. The customization of the model to each user follows the same process defined when validating the QCD intervals [Sanz et al. 2015].

3 QCHarm

QCharm is defined for QCNAME2 to QCNAME5. The gray-scale colors defined in QCNAME1 are not included, because previous research shows they can harmoniously combine with all the rest of colors.

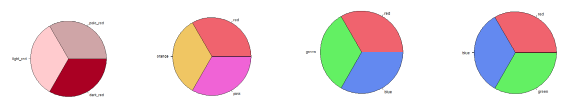
The basic, and most common, color combination operators, namely the *monochromatic, analogous, triad, split-complementary* operations have been defined. Except for the case of the monochromatic operation, the other three operations are defined for each QCNAME independently. Thus, it is necessary to define a fifth operation, the *combining operation*, which will allow the creation palettes of three colors belonging to different sets (QCNAME).

To formalize the QCharm color operations, colors in the CND are represented by an index based on color neighborhood, as follows: 0: green, 1: turquoise/cyan, 2: blue, 3: purple, 4: pink/magenta, 5:red, 6:orange, and 7:yellow. Note that the relationships are circular. Moreover, as it is possible to have four adjectives, following the same process we give a number to each adjective as follows: 0:light, 1:without adjective, 2: dark, and 3:pale. Thus, next sections defines the 5 operations given the set of hues as H={0,1,2,3,4,5,6,7}, and the set of adjectives as A={0,1,2,3}.

Monochromatic color combination is achieved by using only one color, but applying it in different shades and tints. In QCharm these palettes are defined as the combination of a color name in any set, and the corresponding colors in 2 diferent sets (the sets are QCNAME2, QCNAME3 , QCNAME4, or QCNAME5) (Figure 4a). Formally, this means:

$$f_M(h,a) = \{(h,a+1),(h,a+2),(h,a+3)\}(mod8, mod4)$$

where h∈H and a∈A



(a) (b) (c) (d)

Figure 4. Examples of QCharm color combinations based on the *red* color (a) monochromatic; (b) analogous; (c) triadic and (d) split-complementary.

Analogous color palettes use colors that are next to each other on the color wheel (Figure 4b). Formally:

$$f_1(h,a) = \{(h-1,a),(h,a),(h+1,a)\}(mod8, mod4)$$

where h∈H and a∈A

A triadic color scheme uses colors that are evenly spaced around the color wheel (Figure 4c). Formally,

$$f_2(h,a) = \left\{\left(h - round\left(\frac{length(H)}{3}\right), a\right), (h,a), (h + round(\frac{length(H)}{3}\right\}(mod8, mod4)$$

where h∈H and a∈A

To define the **split-complementary operation** it is first necessary to define the complementary operation. Colors that are opposite each other on the color wheel are considered to be complementary colors. The *split-complementary* operation is a variation of the complementary color scheme. In addition to the base color, it uses the two colors adjacent to its complement (Figure 4d).

The palette combination operation can be seen as a generalisation of the other operations as (Figure 5):

$$f_{kR}(h,a) = \{h-k, a(R)), (h, a(R)), (h + k, a(R))\}(mod8, mod4)$$

where h∈H and a∈A, k∈{1,2,3} and a(R) is a number between {0,1,2,3}

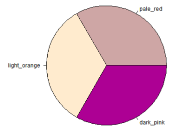


Figure 5. An example of a palette variation in QCharm, based on the analogous palette for the red color.

To demonstrate the possibilities offered by the QCharm theory, it has been implemented as a module in the R programming language, and used it to create a simple application intended to help designers to find harmonious palettes of colors. The user interface guides the designer through three steps. First, the designer is asked to pick a color on which she desires to base the palette (Figure 6). Then, the designer is presented a set of palettes constructed using the basic operations described in this paper (Figure 7). Finally, the application finds all possible palettes variations that can be computed based on the basic harmonic palette chosen by the user. The palettes are scored using a quality model and presented in descending order of score (Figure 8).

In order to obtain a ranking of all the palettes, a learned regression model that tried to learn user preferences based on the ColorLovers dataset has been developed [Museros et al. 2016]. The score is computed by averaging the user-provided scores for each palette. It is a relatively large dataset that contains scores for 174882 palettes (Figure 8).

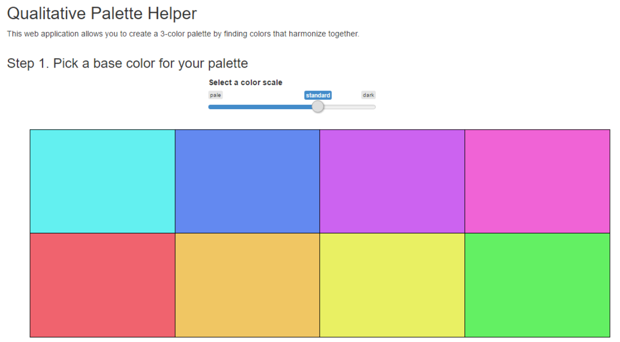


Figure 6. Choosing a base color for a palette.

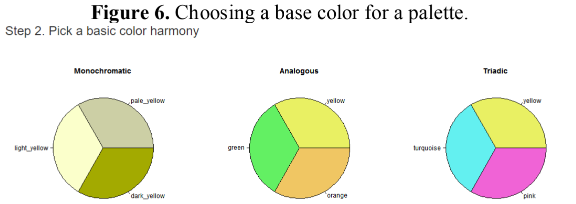


Figure 7. Basic harmonic palettes (in this example, based on yellow).

Step 3. Pick a palette

The following table shows all variations of the colours in your chosen harmony, sorted by a model of preference.

Show 10 entries Search:

	color1	color2	color3	score
1	pale_yellow	light_green	light_orange	3.181257
2	pale_yellow	pale_green	light_orange	3.168612
3	light_yellow	light_green	light_orange	3.129953
4	pale_yellow	light_green	pale_orange	3.122678
5	light_yellow	pale_green	light_orange	3.117308
6	pale_yellow	light_green	orange	3.116173
7	pale_yellow	pale_green	pale_orange	3.110033
8	pale_yellow	pale_green	orange	3.103528
9	light_yellow	light_green	pale_orange	3.071374
10	light_yellow	light_green	orange	3.064869

Showing 1 to 10 of 64 entries Previous 1 2 3 4 5 6 7 Next

Figure 8. Variations of the colors in the palette chosen by the user. The score is obtained by a learned preference model.

4 QArt-Learning: a Qualitative Model for Categorizing Paintings in Art Styles

Qart-Learning [Falomir et al. 2018] is constructed using K-Nearest Neighbour (K-NN) and Support Vector Machine (SVM) supervised learning [Gonzalez-Abril et al. 2009b]. These learning models are based on similarity functions. Therefore, given an image its qualitative histogram is calculated, and the qualitative histogram of each image is an input object of the learning algorithm. The similarity measure t between the histograms is calculated in function of the similarity between two name colors of QCHarm. This similarity is calculated using the CND constructed for the model (Figure 1b). In fact, in order to calculate the similarity between two color names, an interval distance has been used [Falomir et al. 2018]. Also, some other global quantitative features have been used as parameters of the learning algorithm, which are: the arithmetic average of the hue coordinate; the circular average of the hue coordinate; arithmetic average of the lightness coordinate; arithmetic average of the saturation coordinate in the HSL cylinder representation; arithmetic average of the saturation coordinate in the HSL bicone representation; average brightness; logarithmic average brightness; the brightness contrast.

In order to test the QArt-Learn approach, three painting styles were selected: Baroque, Impressionism and Post-Impressionism (Figure 9). Two representative authors were selected for each style, Velázquez and Vermeer as Baroque painters, Monet and Renoir as Impressionists, and van Gogh and Gauguin as Post-Impressionists.

The dataset selected for the experimentation was the Painting-9110 dataset by Khan et al. [2014]. From this dataset, the paintings by the selected authors were extracted, which resulted in 252 images: 74 for Baroque style, 85 as Impressionist paintings and 93 as Post-Impressionist paintings.

Figure 9. Paintings corresponding to the Baroque style (authored by Velázquez (a and b) and Vermeer (c)), Impressionism style (authored by Monet (d and e) and Renoir (f)) and Post-Impressionism style (authored by van Gogh (g and h) and Gauguin(i)) are shown.

Those images were normalized to 10.000 pixels. For each painting image, their QCD color features were extracted.

In the K-NN experimentation, the number of considered nearest neighbors was k = 2, … , 20. The number of training vectors, denoted by *NTraining*, was randomly chosen, where *NTraining*={90+u·15,for u=0,1,…,6},and the rest of instances were the set of test vectors. This procedure was repeated 100 times in order to ensure good statistical behavior.

The accuracy performance for the 1-versus-rest SVM was evaluated on models using a Radial Basis Function (RBF) kernel with σ (RBF width) and C (regularization term) initially explored on a two-dimensional grid, $\sigma = [2^{-4},2^{-3},...,2^{5},2^{6}]$ and $C = [2^{-2},2^{-1},...,2^{6},2^{7}]$, after a finer grid is chosen. The criteria selected to estimate the generalized accuracy was 10-fold cross-validation on the whole set of training data. Also, this procedure is repeated 100 times.

The results obtained by the k-NN are given in Table 1. The higher accuracy for each set of features is highlighted by a grey background and the best result for each *NTraining* time is indicated in bold. Let us indicate that the accuracy results are similar in- dependently of the number of neighbors used (k). Genrally, the higher the training *(NTraining*), the higher the accuracy, for any training set. Note also that, by using 180 training sets, a higher accuracy is obtained, but it only differs in 3 points from that obtained by 90 training sets. Thus, as training further is not efficient, the QArt-Learn approach stops. It is worth noting that, for any *NTraining*, the obtained accuracy is always higher when considering only the 37-color-palette (QCD) than when using the eight average global parameters. Sometimes the results are better when considering only the QCD than when considering all the features together.

	All features		37-QC-palette		Average features	
NTraining	K	%	K	%	K	%
180	6	**67.27**	8	66.69	6	65.86
165	6	**66.70**	5	65.79	4	65.57
150	6	65.31	11	**66.14**	6	64.71
135	4	64.70	9	**65.16**	4	64.45
120	4	64.43	6	**64.93**	4	63.79
105	4	63.94	8	**64.74**	4	62.75
90	9	63.93	7	**63.98**	4	62.13

Table 1: Accuracy values obtained by k-NN using the QCD and/or the 8 average color features.

The results obtained by the SVMs using the best cross-validation mean rate are shown in Table 2. The higher accuracy for each set of features is highlighted by a grey background and the best result for each NTraining time is indicated in bold. Again, analyzing the results for any *NTraining*, the accuracy is always higher when using only the QCD than when all features or the eight average global parameters are considered.

	All features		37-QC-palette		Average features	
NTraining	(σ,C)	%	(σ,C)	%	(σ,C)	%
180	(1.68, 1.19)	64.87	(1.68, 1.00)	**65.34**	(0.21, 1.00)	63.02
165	(2.00, 1.00)	64.28	(1.68, 2.00)	**64.82**	(0.25, 1.00)	62.84
150	(2.00, 1.00)	64.55	(1.69, 1.41)	**65.13**	(0.30, 0.84)	62.37
135	(1.68, 1.41)	63.57	(2.00, 0.71)	**65.56**	(0.30, 1.00)	62.29
120	(2.00, 0.84)	63.52	(2.00, 1.19)	**64.91**	(0.35, 0.70)	62.61
105	(2.00, 0.84)	63.07	(2.00, 0.84)	**64.59**	(0.29, 1.00)	62.38
90	(1.68, 1.68)	62.80	(1.68, 1.00)	**64.18**	(0.35, 0.84)	61.51

Table 2: Accuracy values obtained by SVMs using the QCD and/or the 8 average color features.

Both classifiers, k-NN and SVM, performed with an accuracy over 65% using the QCD which shows that there is a qualitative color palette that can describe each style. The eight quantitative global average features are obtaining the lower accuracy results for both classifiers k-NN and SVM. This shows that both the QCD and the global quantitative average features are quite equivalent for this classification problem. However, the QCD, as it is based on color names, can be used to give a definition of the style in color names or to provide an explanation of the classification outliers, while the quantitative features do not have a correspondent alignment with linguistic concepts.

Analizing the confussion matrix in Table 3, it can be concluded that paintings belonging to Impressionism and Post-Impressionism styles are more difficult to distinguish, as it happens for the human eye, by only taking into account color features.

	Real	**Prediction**		**Total**	**Percentage of success**	
		Impressionism/Post-Imp.	Baroque	7200	Impressionism/Post-Imp.	Baroque
All features	Impressionism/Post-Imp.	4727	313	5040	93.79	06.21
	Barroque	513	1647	2160	23.75	76.25
37-QC	Impressionism/Post-Imp.	4780	305	5085	94.00	6.00
	Baroque	537	1568	2105	25.51	74.49
Average features	Impressionism/Post-Imp.	4943	219	5162	95.76	4.24
	Baroque	580	1458	2038	28.46	71.57

Table 3: Confusion matrix for NTraining = 180 using 6-NN and the 37-QC-palette.

4 Conclusions

A validated QCD model is presented in this paper. The necessity to customize it to a user profile has also been demonstrated, and a method to carry out this customization has been provided. Also a Qualitative Harmony theory (QCharm) has been presented. QCHarm is able to define harmonic palettes based on a color label by applying a set of operations. We have shown its applicability by building a prototype palette recommendation application. QArt-Learn approach based on the QCD is also presented, which automatize the color categorization of paintings in three art styles: Baroque, Impressionism and Post- Impressionism. Two machine learning classifiers, K-NN and SVM were built and tested. The results of this experimentation show that both classifiers, k-NN and SVM, performed with an accuracy over 65%, showing that a color palette may describe a painting style. The quantitative global average features did not improve the accuracy results obtained significantly. Let us highlight that, using the color names used by the QArt-Learn approach, an artificial agent could describe a painting color style to a non-expert human.

As future work, we will intend to extend the QArt-learn to more painters and styles, and to provide an explanation of the classification outliers using also the colors appearing or not appearing in the paintings. And finally, we are also interested in not only rating the color palettes generated but also associate an emotion to them.

Acknowledgments

Dr. Museros, Dr. Sanz and Prof. Dr. Gonzalez-Abril acknowledge the funding by the Spanish Ministry of Economy and Competitiveness (TIN2017-88805-R) and Universitat Jaume I (UJI-B2017-73). Dr.-Ing. Falomir acknowledges the funding and support by the Bremen Spatial Cognition Center (BSCC), and the University Bremen under project Cognitive Qualitative Descriptions and Applications (CogQDA).

References

[Abelson *et al.*, 1985] Harold Abelson, Gerald Jay Sussman, and Julie Sussman. *Structure and Interpretation of Computer Programs.* MIT Press, Cambridge, Massachusetts, 1985.

[Cohn and Renz 2007] Cohn, A.G., Renz, J. *Qualitative Spatial Reasoning, Handbook of Knowledge Representation*. Elsevier, Wiley-ISTE, London, 2007.

[Conway 1992] D. Conway, An experimental comparison of three natural language colour naming models. In: *Proc. East–West International Conference on Human–Computer Interactions*, pp. 328–339, 1992

[Corridoni et al. 1998] J.M. Corridoni, A.D. Bimbo, E. Vicario, Image retrieval by color semantics with incomplete knowledge, *Journal of the American Society for Information Science 49* (3) (1998) 267–282.

[Falomir et al. 2013] Z. Falomir, Ll. Museros, L. González-Abril, I. Sanz, A model for qualitative colour comparison using interval distances, *Displays, 34 (4)*: 250-257, 2013.

[Falomir et al. 2018] Falomir, Z., Museros, Z., Sanz, I., Gonzalez-Abril, L., Categorizing paintings in art styles based on qualitative color descriptors, quantitative global features and machine learning (QArt-Learn). *Expert Syst. Appl. 97*: 83-94 (2018).

[Freksa 2013] C. Freksa, Spatial computing – how spatial structures replace computational effort. In: M. Raubal, D. Mark, A. F. (Eds.), *Cognitive and linguistic aspects of geographic space*, Springer, Heidelberg, 2013.

[Gatys et al. 2015] Gatys, L.A., Ecker, A.S., Bethge, M. *A neural algorithm of artistic style*. URL: http://arxiv.org/abs/1508.06576, 2015

[Gonzalez-Abril et al. 2009] L. Gonzalez-Abril, F.J. Cuberos, F. Velasco y J.A. Ortega. Ameva: An autonomous discretization, *Expert Systems with Applications Volumen 36, Issue 3, Part 1*: 5327-5332, 2009.

[Gonzalez-Abril et al. 2009b] Gonzalez-Abril, L., Velasco, F., Ortega, J.A., Cuberos, F.J., 2009. A new approach to qualitative learning in time series. *Expert Systems with Applications 36*, 9924–9927.

[Itten 1970] Itten. In: Birren F, editor. *The Elements of Color*, translated by Enrnst van Hagen. New York: Van Nostrand-Reinhold; 1970.

[Jahanian et al. 2013] Jahanian, A., Liu, J., Lin, Q., Tretter, D., O'Brien-Strain, E., Lee, S., Lyons, N., and Allebach, J. Recommendation System for Automatic Design of Magazine Covers. *Proc. IUI'13*, pp. 95- 105, March 10-22, 2013.

[Kay and Regier, 2006] P. Kay and T. Regier. Language, thought, and color: Recent developments. *Trends in Cognitive Sciences*, 10(2): 51–54, 2006.

[Karayev et al. 2013] Karayev, S., Hertzmann, A., Winnemoeller, H., Agarwala, A., Darrell, T. Recognizing image style. *CoRR abs/1311.3715*. URL: http://arxiv.org/abs/1311.3715, 2013

[Khan et al. 2014] Khan, F.S., Beigpour, S., van de Weijer, J., Felsberg, M., 2014. Painting-91: a large scale database for computational painting categorization. *Machine Vision and Applications 25*, 1385–1397. doi:10.1007/s00138-014-0621-6.

[Kuang et al. 2014] Kuang, Z., Lu, P., Xiaojie, W., Xiaofeng L. Learning self-adaptative color harmony model for aesthetic quality classification. *Proc. SPIE 9443, Sixth International Conference on Graohic and ImageProcessing,* 944310, 2014.

[Lammens 1994] J.M.G. Lammens, A computational model of color perception and color naming, Ph.D. thesis, Faculty of the Graduate School of State University of New York at Buffalo, USA, 1994.

[Li and Chen 2009] Li, C., Chen, T. Aesthetic visual quality assessment of paintings. Selected Topics in Signal Processing*, IEEE Journal of 3*, 236–252, 2009

[Li et al. 2012] Li, J., Yao, L., Hendriks, E., Wang, J.Z. Rhythmic brushstrokes distinguish van gogh from his contemporaries: Findings via automated brushstroke extraction. *IEEE Transactions on Pattern Analysis and Machine Intelligence 34*, 1159–1176. doi:10.1109/TPAMI. 2011.203, 2012.

[Liu et al. 2004] Y. Liu, D. Zhang, G. Lu, W. Ma, Region-based image retrieval with perceptual colors, in: *Proc. of Pacific-Rim Multimedia Conference (PCM2004)*, 2004, pp. 931–938.

[Menegaz et al., 2007]] G. Menegaz, A.L. Troter, J. Sequeira, J.M. Boi, A discrete model for color naming, *EURASIP Journal of Applied Signal Process, Special Issue on Image Perception* (2007) pp. 1–10.

[Mensink and Gemert 2014] Mensink, T., van Gemert, J. The rijksmuseum challenge: Museum-centered visual recognition, in: *ACM International Conference on Multimedia Retrieval* (ICMR), 2014

[Mojsilovic 2005] A. Mojsilovic, A computational model for color naming and describing color composition of images, *IEEE Transactions on Image Processing 14 (5) (2005)* pp. 690–699.

[Moretti et al. 2013] Moretti, G., Lyons, P., and Marsland, S. Computational Production of Color Harmony. Part 1: A Prototype Color Harmonization Tool. *Color Research & Applications 38(3)*, 2013

[Munsell 1921] Munsell, AH. *A Grammar of Color*. New York: Van Nostrand-Reinhold, 1921.

[Museros et al. 2016] Museros, Ll., Sanz, I., Falomir, Z., Gonzalez-Abril, L. A Qualitative Color Harmony Theory. *Frontiers in Artificial Intelligence and Applications Volume 288: Artificial Intelligence* Research *and Development*, pp. 98-107, 2016

[Nishiyama et al. 2011] Nishiyama, M., Okabe, T., Sato, I., Sato, Y. Aesthetic Quality Classification of Photographs Based on Color Harmony. *Computer Vision and Pattern Recognition*, 33-40, 2011.

[O'Connor 2014] O'Connor, Z. *Color Combination Techniques*. Sydney AUS: Design Research Associates, 2014.

[O'Donnovan et al. 2011] O'Donovan, P., Agarwala, A., and Hertzmann, A. Color compatibility From Large Datasets. *Proc. SIGGRAPH'11 ACM, article no. 63*, 2011.

[Orlov et al. 2007] Orlov, N., Johnston, J., Macura, T., Shamir, L., Goldberg, I. Computer vision for microscopy applications, in: Obinata, G., Eds, A.D. (Eds.), *Vision Systems – Segmentation and Pattern Recognition.* ARS Pub., Viena, Austria, pp. 221–242, 2007.

[Ou and Ronnier Lou 2006] Ou, L., Ronnier Lou, M. A Color Harmony Model for Two-Color Combinations. *Color Research and Application, Volume 31(3)*, 191-204, 2006

[Ou et al. 2012] Ou, L., Ronnier Lou, M., Hu, N., Chen, H., Guan, S., Woodcock, A., Caivano, J.L., Huertas, R., Treméau, A., Billger, M., Izadan, H., and Richter, K. A Cross-Cultural Comparison of Color Emotion for Two-Color Combinations, *Color Research and Applications, Volume 37(1)*, 23-43, 2012.

[Palmer and Griscom 2013] Palmer, S., Griscom, W. Accounting for taste: Individual differences in preference for harmony. *Psychon Bull Rev 20*:453-461, 2013. DOI: 10.3758/s13423-012-0355-2

[Sanz et al. 2015] I.Sanz, Ll. Museros Cabedo, Z. Falomir, and L. González-Abril: Customising a qualitative colour description for adaptability and usability. *Pattern Recognition Letters 2-10* (2015)

[Schloss and Palmer 2011] Schloss, K., and Palmer, E. Aesthetic response to color combinations: preference, harmony, and similarity. *Atten Percept Psychophys 73*:551–571, 2011

[Seaborn et al. 2005] M. Seaborn, L. Hepplewhite, T.J. Stonham, Fuzzy colour category map for the measurement of colour similarity and dissimilarity, *Pattern Recognition 38 (2)* (2005) 165–177.

[Shamir et al. 2010] Shamir, L., Macura, T., Orlov, N., Eckley, D.M., Goldberg, I.G. Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. *ACM Trans. Appl. Percept. 7, 8*:1–8:17. doi:10. 1145/1670671.1670672, 2010

[Shamir et al. 2012] Shamir, L., Tarakhovsky, J.A.Computer analysis of art. *J. Comput. Cult. Herit. 5, 7*:1–7:11. doi:10. 1145/2307723.2307726, 2012

[Shen 2009] Shen, J. Stochastic modeling western paintings for effective classification. *Pattern Recognition 42*, 293 – 301. doi:10.1016/j.patcog.2008.04.016, 2009

[Siddiquie et al. 2009] Siddiquie, B., Vitaladevuni, S., Davis, L. Combining multiple kernels for efficient image classification, in: *Applications of Computer Vision (WACV), 2009 Workshop on*, pp. 1–8. doi:10.1109/WACV.2009. 5403040, 2009.

[Soto-Hidalgo et al. 2010] J.M. Soto-Hidalgo, J. Chamorro-Martínez, D. Sánchez, A new approach for defining a fuzzy color space. In:*Proc. 2010 IEEE International Conference on Fuzzy Sustems (FUZZ)*, pp.1-6, 2010.

[Stanchev et al. 2003] P.L. Stanchev, D. Green Jr., B. Dimitrov, High level colour similarity retrieval, *International Journal on Information Theories and Applications 10* (3) (2003) pp. 363–369.

[Yelizaveta et al. 2005] Yelizaveta, M., Tat-Seng, C., Irina, A. Analysis and retrieval of paintings using artistic color concepts, in: Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on, pp. 1246–1249. doi:10. 1109/ICME.2005.1521654, 2005.

Hesitancy and consensus measures to understand ratings: An application to hotel recommendations

Nguyen J., Montserrat-Adell J., Agell N., Sánchez M. and Ruiz F.J.

Abstract

When searching for recommendations online, users are often presented with similar items all having the same ratings. It is left to the user to discern which rated item is best from other information. In order to facilitate this process we propose a new methodology to associate a measure of consensus to the ratings derived from a group of reviewers. The opinion of reviewers are considered not only from rating values but from written reviews. This allows us to express reviewer opinions as interval ratings by means of hesitant fuzzy linguistic term sets and compute a consensus. A real case example is provided to show the potential of the proposed methodology. The example provided is based on 966 TripAdvisor reviews of the twenty most reviewed hotels during 2017 in Rome at the time the data was taken. The consensus measure allows us to better understand hotel reviews.

1 Introduction

Marketing research has found that consumers influence each other in their decision making process [Dichter, 1966]. On internet platforms, this influence is derived from ratings and reviews [Amblee and Bui, 2011]. These reviews facilitate the decision-making process between people using the same platform. According to Neilson[1], 70% of social media users go online to read about other people's experiences with an item at least once a month.

With the abundant amount of information available to consumers on the Internet, recommender systems assist with the information overload by limiting information according to a consumer's interest [Adomavicius and Tuzhilin, 2005]. Recently, more review-based recommender systems have been developed with the availability of user-generated reviews. Advanced text analytics and opinion-mining enable the extraction of key features from these reviews such as reviewer emotions [Chen *et al.*, 2015]. Traditionally, recommender systems utilize information from ratings or text to rank items according to user preferences. However, the ranked items may all be rated with the same value. For example, a travel site may list the top 10 hotels in a city but all of them are rated 5 stars appearing uniform to the online user.

One study by [Gavilan *et al.*, 2018] found that users trust low ratings more than high ratings. In addition, they identified a moderating effect between the relationship of the number of ratings and their trustworthiness. The trustworthiness of low ratings was not impacted by the number of ratings. In contrast, high ratings were found to be trustworthy only when expressed by a high number of reviews.

Several studies have shown that people are more comfortable expressing their preferences in an abstract manner based on linguistic models rather than purely in a quantitative manner [Alonso *et al.*, 2010; Herrera and Herrera-Viedma, 2000; Travé-Massuyès *et al.*, 2005]. Decision-support systems which consider linguistic values to describe alternatives have been developed to facilitate Group Decision-Making (GDM). These linguistic descriptors enable systems to handle the imprecision involved in decision processes as intervals or fuzzy values [Alonso *et al.*, 2010; Montes *et al.*, 2015]. Rodríguez et al. [Rodríguez *et al.*, 2012] introduced Hesitant Fuzzy Linguistic Term Sets (HFLTSs) over a well-ordered set of linguistic labels to reflect the hesitancy inherent in human reasoning. Several decision making approaches and applications based on HFLTS have been developed [Fahmi *et al.*, 2016; Rodríguez *et al.*, 2013; Rodríguez *et al.*, 2014; M. *et al.*, 2016]. In addition, some contributions have analyzed the quantification of the level of agreement or consensus among reviewers by means of HFLTS [Dong *et al.*, 2015; Rodríguez and Martínez, 2015; Wu and Xu, 2016b; 2016a].

In spite of a recommender system's ability to narrow information specific to a user's interest, users are still left with the task of differentiating between the suggested items. This task is still costly to the consumer in time and energy. Therefore, we propose a methodology to ease this differentiating process by introducing a measure of consensus representing the agreement among reviewers of an individual item. This paper moves in two directions: first, building hesitant terms from reviewer ratings and written reviews and second, to measure the consensus of these reviews for a single item to discriminate between items in the same rating category.

In this paper we consider HFLTS to jointly represent the ratings and the text for each review. This new rating is ob-

[1] http://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2012-Reports/The-Social-Media-Report-2012.pdf

tained by incorporating the values obtained from sentiment analysis of the written reviews with the reviewer's rating. Then, a measure of consensus defined in [Montserrat-Adell *et al.*, 2018] that takes into account both the agreement and the disagreement among reviewers is taken. In this way, the methodology allows us to distinguish differences among ratings that were initially equally informed.

The rest of this paper is structured as follows: firstly, Section 2 summarizes the basic concepts already presented in a previous study conducted by [Montserrat-Adell *et al.*, 2018]. Section 3 introduces our proposed methodology and a real case application of the proposed methodology. Finally, Section 4 contains the main conclusions and lines of future research.

2 Preliminaries

This section presents a summary of basic concepts of HFLTS, including the distance between HFLTS and the measure of consensus that will be used in the proposed methodology.

From here on, let $\mathcal{S}$ denote a finite totally ordered set of linguistic terms, $\mathcal{S} = \{a_1, \ldots, a_n\}$ with $a_1 < \cdots < a_n$. For the rest of this article, $\{x \in \mathcal{S} \mid a_i \leq x \leq a_j\}$ is denoted as $[a_i, a_j]$ if $i < j$ or $\{a_i\}$ if $j = i$. According to [Montserrat-Adell *et al.*, 2016], $\mathcal{H}_{\mathcal{S}}$ is defined as the set of all possible HFLTS over $\mathcal{S}$ including the empty HFLTS, $\{0\}$. The set $\mathcal{H}_{\mathcal{S}}$ is extended to $\overline{\mathcal{H}_{\mathcal{S}}}$, to include the concept of *negative HFLTS* as gaps between pairs of HFLTS.

In addition, in the frame of $\overline{\mathcal{H}_{\mathcal{S}}}$, an *extended inclusion relation* is presented in [Montserrat-Adell *et al.*, 2016], and in this context we consider the extended connected union and extended intersection operators.

1. The *extended intersection* of H_1 and H_2, $H_1 \sqcap H_2$, is the largest element in $\overline{\mathcal{H}_{\mathcal{S}}}$ that is contained in H_1 and H_2 according to the extended inclusion relation.
2. The *extended connected union* of H_1 and H_2, $H_1 \sqcup H_2$, is the least element in $\overline{\mathcal{H}_{\mathcal{S}}}$ that contains H_1 and H_2 according to the extended inclusion relation.

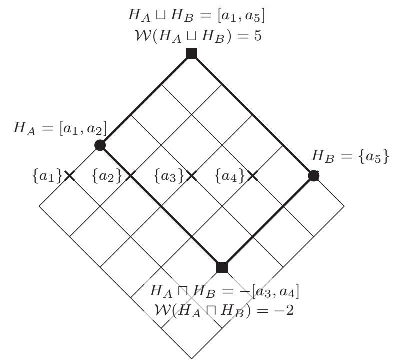
Finally, we consider the distance between two HFLTS as defined in ([Montserrat-Adell *et al.*, 2016]). Given H_1 and $H_2 \in \overline{\mathcal{H}_{\mathcal{S}}}$, the *width* of H, $\mathcal{W}(H)$, is defined as the cardinality of H if $H \in \mathcal{H}_{\mathcal{S}}$ or $-card(-H)$ if H is a negative HFLTS. Then the distance between HFLTS in $\overline{\mathcal{H}_{\mathcal{S}}}$ is computed from $H_1, H_2 \in \overline{\mathcal{H}_{\mathcal{S}}}$, as:

$$D(H_1, H_2) := \mathcal{W}(H_1 \sqcup H_2) - \mathcal{W}(H_1 \sqcap H_2). \qquad (1)$$

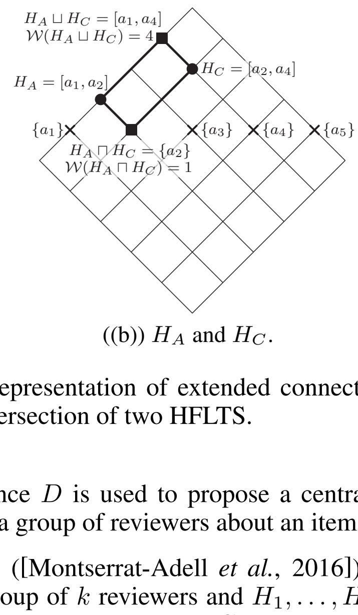
The following example illustrates the previous concepts:

Example 1 *Given a set of possible traveler ratings in linguistic terms:* $\mathcal{S} = \{a_1, a_2, a_3, a_4, a_5\}$, *being* a_1 = *terrible,* a_2 = *poor,* a_3 = *average,* a_4 = *very good and* a_5 = *excellent, three travelers provided the following linguistic assessments of a hotel:* A = *"really bad",* B = *"excellent" and* C = *"OK" and their corresponding HFLTS by means of* $\mathcal{S}$ *are* $H_A = [a_1, a_2]$, $H_B = \{a_5\}$ *and* $H_C = [a_2, a_4]$ *respectively. Figure 1 shows the extended connected union and extended intersection of* H_A *and* H_B *as well as* H_A *and* H_C.

According to these results, $D(H_A, H_B) = 5 - (-2) = 7$ and $D(H_A, H_C) = 4 - (1) = 3$.



((a)) H_A and H_B.



((b)) H_A and H_C.

Figure 1: Representation of extended connected union and extended intersection of two HFLTS.

The distance D is used to propose a central opinion (or centroid) of a group of reviewers about an item λ as follows:

Definition 1 ([Montserrat-Adell *et al.*, 2016]) Let λ be an item, G a group of k reviewers and $H_1, \ldots, H_k$ the HFLTS of λ provided by the reviewers in G. Then, the *centroid of the group* is:

$$C_o = arg \min_{H \in (\mathcal{H}_{\mathcal{S}}^*)} \sum_{i=1}^{k} D(H, H_i). \qquad (2)$$

The centroid is similiar to the median of a group. It is a central measure for ordinal scales with hesitancy. In order to ease the calculation of the centroid, [Montserrat-Adell *et al.*, 2016] proved that, for each specific alternative $\lambda \in \Lambda$, if $F_H^p(\lambda) = [a_{i_p}, a_{j_p}]$ is the HFLTS used by DM p to assess λ, then the set of all the HFLTS associated to the centroid of the group for λ is:

$$\{[a_i, a_j] \in \mathcal{H}_{\mathcal{S}}^* \mid i \in \mathcal{M}(i_1, \ldots, i_k), j \in \mathcal{M}(j_1, \ldots, j_k)\}, \qquad (3)$$

where $\mathcal{H}_{\mathcal{S}}^* = \mathcal{H}_{\mathcal{S}} - \{0\}$ and $\mathcal{M}(\ )$ is the set that contains just the median of the values if k is odd or any integer number between the two central values sorted from smallest to largest if k is even.

Example 2 *Let G be a group of 5 reviewers who are assessing a hotel λ by means of HFLTS over the set $\mathcal{S}$ from Example 1, and let H_1, H_2, H_3, H_4, H_5 be the HFLTS describing their corresponding assessments shown in Table 1. Then, the centroid of the group, C_o, can be calculated through the medians as seen in the same table. Figure 2 is a representation of the centroid, C_o with respect to the HFLTS.*

	H_1	H_2	H_3	H_4	H_5	$\boldsymbol{C_o}$
λ_1	$[a_2, a_3]$	$\{a_2\}$	$[a_4, a_5]$	$[a_1, a_2]$	$[a_1, a_4]$	$\boldsymbol{[a_2, a_3]}$

Table 1: Centroid of the group G for λ.

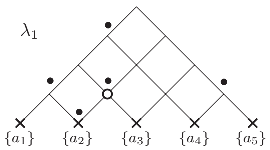


Figure 2: H_1, H_2, H_3, H_4, H_5 and C_o from Example 2.

Next, in this section we introduce the consensus degree defined in [Montserrat-Adell *et al.*, 2018] that seeks to quantify the agreement between a group of reviewers when rating an item.

Definition 2 Let G be a group of k reviewers of an item λ, and $H_1, \ldots, H_k$ be their respective ratings by means of HFLTS. Let C_o be the centroidal review of the group. Then, the *degree of consensus of G on λ* is defined as:

$$\delta_\lambda(G) = 1 - \frac{\sum_{i=1}^{k} D(C_o, H_i)}{k \cdot (n-1)}. \tag{4}$$

Note that $0 \leq \delta_\lambda(G) \leq 1$ due to $k \cdot (n-1)$ is an upper bound of the addition of distances between the centroid and the HFLTS reviewers [Montserrat-Adell *et al.*, 2018].

Example 3 Following Example 2, G is a group of 5 reviewers who are assessing a hotel λ by means of HFLTS over the set $\mathcal{S}$. In Table 2, D_i are the distances from each assessment to the central opinion and $\delta_\lambda(G)$ the degree of consensus of G.

	D_1	D_2	D_3	D_4	D_5	$\sum_{i=1}^{5} D_i$	$\delta_\lambda(G)$
λ	0	1	4	2	2	9	0.45

Table 2: Consensus on the evaluation of a hotel

3 A real case example to improve recommendations

Recommender systems rely on a set of ratings for any particular item in order to provide users with a ranked list of items. Reviewers may provide ratings in different formats such as numerical ratings, number of stars, or written reviews. Linguistic ratings may be associated with the number of stars. For example, on TripAdvisor, an "average" rating is equivalent to three stars on a scale from one to five. We propose that written reviews along with ratings can be used to determine more representative linguistic expressions of human assessments of an item. In other words, from each rating and written review we propose to generate a HFLTS. If we consider an item to be a hotel, our methodology follows the steps of Figure 3. First, for each hotel and each reviewer we generate a HFLTS. Next, we compute the centroid of each hotel from all the reviewers' opinions given by HFLTS. Third, we measure the degree of consensus for each hotel via the distance between the reviews and the centroid. Then, a new rating for each hotel can be expressed as a combination of the centroid and consensus measure. Given this new rating, we are able to "totally" order the hotel reviews for each category of rating.

3.1 Data Set

In order to demonstrate our methodology on a real case, we selected TripAdvisor reviews of hotels in Rome. Xiang et al. [Xiang *et al.*, 2017] found TripAdvisor reviews to have higher overall quality when compared to other online sites. In addition, the authors concluded that the connections between ratings, helpfulness, and review topics are stronger in TripAdvisor reviews in comparison to other sites suggesting some consistency between written reviews and ratings.

We used a data set of TripAdvisor bed and breakfast reviews in Italy between 2002 and 2017 provided on Kaggle [2]. The initial set of data contained 223,089 reviews for 3,760 hotels. For each review the title, date, rating, text, language, reviewer id, and property id were provided. In addition, each review contains a rating value considered as an ordinal scale $R \in \{1, \ldots, 5\}$. For each hotel the hotel id, name, total number of reviews, average displayed on TripAdvisor, address, and coordinates were provided. We began by narrowing down our data set. We focused our data set to those reviews written in 2017. Selected only those properties with an address in Rome. Then, eliminated all non-English reviews. The 2017 dataset contained 2,715 hotels and 31,396 reviews of which 36% were written in English.

We removed any duplicated reviews and selected the twenty most reviewed hotels. The final data set contained 966 reviews. We processed the data and reviewed descriptive statistics including average text length, number of reviews per hotel, percentage of English reviews following [Xiang *et al.*, 2017].

3.2 Experimental Approach

All reviews in English were pre-processed in preparation for semantic analysis. Common text transformations were applied removing numbers and non-alphabetical characters. Then, stop words that do not contribute to the review meanings were removed. For each hotel we applied the steps in Section 3.

[2] `https://www.kaggle.com/nicodds/rome-b-and-bs`

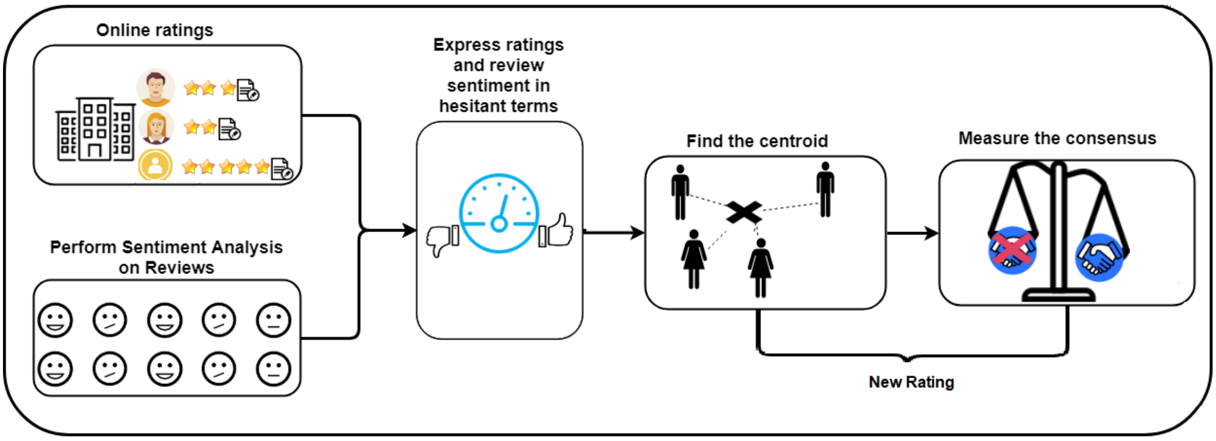


Figure 3: Methodology to combine ratings and reviews into HFLTS and express their consensus

In order to use a more representative linguistic expression, we first need to determine the sentiment of each review. We propose to apply sentiment analysis with the AFINN lexicon [Nielsen, 2011] to evaluate the opinions in the text. AFINN is a list of 2477 English words and phrases rated for valence with an integer between minus five and plus five. Here minus is an indicator of "badness" and plus of "goodness" of a review. In a study of twenty-four off-the-shelf methods of sentiment analysis performed on eighteen labeled datasets, Ribeiro et al. [Ribeiro *et al.*, 2016] found that no single method achieved the best prediction performance across all the datasets tested. The methods were tested on tweet, comment, and review text. AFINN was second in mean rank for three-class classification (positive, negative, and neutral) across the comments datasets behind the VADER method. However, AFINN was eighth for two-class classification (positive and negative) for the reviews datasets ahead of VADER. There was no three-class classification for reviews as the datasets did not contain a considerable number of neutral messages. As our method is interested in modeling all three types of sentiment, we selected AFINN.

Words for each review r are matched to the words in the AFINN dictionary. The output is a set of matching words $\{w_{r1}, \ldots, w_{rp}\}$. For each review r, each word w_j, $j \in \{1, \ldots, p\}$ is associated with a valence $v(w_{rj}) \in \{-5, \ldots, 5\}$ and a frequency occurrence $f(w_{rj}) \in \{1, \ldots, q_r\}$ in the review. An example of the sentiment analysis output after applying AFINN is given in Table 3. The example shows the set of words identified from the AFINN dictionary along with their associated valence for each of the five reviews for hotel #11.

Based on the valence $v(w_{rj})$ and frequency $f(w_{rj})$ of each word in a review, we can assign a HFLTS to each rating. We compute the positive S_r^+ and negative S_r^- sentiment contributions according to Equation 5. Continuing with the example reviews from hotel #11, the resulting positive S_r^+ and negative sentiment S_r^- contributions are provided in Table 4.

Review	Word	Freq	Score
#1	comfortable	1	2
	friendly	1	2
	lovely	1	3
	perfect	1	3
	wonderful	1	4
	worth	1	2
#2	comfortable	1	2
	easy	1	1
	friendly	1	2
	helpful	1	2
	lovely	1	3
#3	recommend	2	2
	superior	2	2
	free	1	1
	lied	1	-2
	mistake	1	-2
#4	clean	1	2
	confusing	1	-2
	fantastic	1	4
	happy	1	3
	helpful	1	2
	laughing	1	1
	love	1	3
	reached	1	1
	safe	1	1
	secured	1	2
	significant	1	1
	stop	1	-1
#5	comfortable	2	2
	cancelled	1	-1
	clean	1	2
	funny	1	4
	overlooked	1	-1

Table 3: Output of sentiment analysis for hotel #11

Review	R	S_r^+	S_r^-
#1	5	1.000	0.000
#2	5	1.000	0.000
#3	1	0.692	0.308
#4	4	0.870	0.130
#5	4	0.833	0.167

Table 4: Review ratings and sentiment contributions for hotel #11

	Review 1	2	3	4	5	
HFLTS H	{5}	{5}	{1}	[3,5]	[3,5]	$C_o = [3,5]$
Distance D	2	2	6	0	0	$\delta_\lambda(G)$ =0.5

Table 5: HFLTS, Distance, centroid, and consensus for hotel #11

$$S_r^+ = \frac{\sum\limits_{\substack{j=1 \\ v(w_{rj})\geq 0}}^{p} v(w_{rj}) * f(w_{rj})}{\sum\limits_{j=1}^{p} |v(w_{rj})| * f(w_{rj})}$$
$$S_r^- = \frac{\sum\limits_{\substack{j=1 \\ v(w_{rj})<0}}^{p} v(w_{rj}) * f(w_{rj})}{\sum\limits_{j=1}^{p} |v(w_{rj})| * f(w_{rj})} \tag{5}$$

Once each review r has been assigned a positive S_r^+ and negative sentiment S_r^- contribution, we propose to define a HFLTS $[H_r^+, H_r^-]$ following Equation 6. We consider a reviewer's hesitancy to be inclusive of the numerical rating R given for the same review r. For each review r in the example for hotel #11, the HFLTS have been determined in Table 5.

$$H_r^+ = min(5, \lfloor R + R * S_r^+ \rfloor)$$
$$H_r^- = max(1, \lfloor R - R * S_r^- \rfloor) \tag{6}$$

In order to determine the centroid C_o of the k reviews for a hotel λ, we apply Definition 1 with H_r being the HFLTS of the reviews as determined in the sample in Table 5. Table 5 shows the computed distances to the centroid for the sample reviews. Next, Definition 2 is applied to measure the degree of consensus of these reviews. Returning our attention to the data set of twenty hotels to which we applied our methodology, the centroid for all the reviews for each hotel is shown in Table 6. As the hotels we selected had too many reviews to show all the values for each, we displayed only a sample. However, the examples from Section 2 illustrate how to complete the computation for the centroid and consesus measure for each hotel.

3.3 Results

We assess our proposed methodology from two aspects. First, we provide the processing time for our proposed method. Second, we evaluate the potential applicability of the proposed method. The real case was implemented on a 2 GHz Intel Core i5 MacBook Pro (2016) with 8 GB of memory. The complete process took 1 minute 53 seconds. In order to test the efficiency of the proposed methodology on a larger data set, we re-ran the real case on the original data set. After removing non-English and duplicate reviews the data set contained 2,719 hotels and 70,571 reviews. This process took 30 hours and 40 minutes. Although the initial set-up took considerable time, future incremental updates with each new review is expected to take little time.

Table 6 is a summary of the reviews and results for the top twenty most reviewed hotels in Rome in 2017. For each hotel, the number of reviews per hotel, average number of words in each written review, and the median rating of the reviews for each hotel is given. HFLTS were defined based on the proposed methodology for each review. The frequency of HFLTS for each hotel is shown along with the computed centroid and consensus of the new ratings for each hotel.

We can compare the results of the proposed HFLTS rating with the median ratings for the TripAdvisor hotels. As can been seen from Table 6, providing both the centroid of the reviews and the degree of consensus among the reviews can potentially offer two benefits. First, the user can discriminate between many hotels having the same rating such as "4 stars" by taking a closer look at the centroids. In these cases, the centroid can tell the user in which direction the reviewer opinions differ from one another. For example, if the centroid is [3,4], the user would know that reviewers of this hotel have some negative impressions different from the "4 star" rating. In contrast, if the centroid were [4,5], the user would know that reviewers tended to have more positive impressions of the hotel. Second, considering the hotels rated with "5 stars", the user can quickly notice that for some hotels the central opinion of reviewers is $\{5\}$. However, taken in combination with the consensus measure, he can differentiate which of those hotels reviewers agree the most deserves "5 stars". The hotels in Table 6 are sorted by the centroid and the consensus measure.

As can be seen from the results prior to applying the proposed methodology, only 19% of the hotels could be differentiated by the ratings alone for the given sample. Following the implementation of the methodology, 99.5% of the hotels could be differentiated.

4 Conclusions and future work

This paper presents a new methodology to associate an interval rating (hesitant term) together with a measure of consensus to the hotel ratings derived from a group of reviewers. Specifically, it gives recommender systems the ability to extend reviewer opinions from ratings to hesitant fuzzy linguistic term sets by combining the opinion of ratings and written reviews. From each set of extended reviewer opinions it considers the centroid to be the global opinion of each hotel. In this way, group consensus can be measured for each hotel and

				Frequency of hesitant terms										
Hotel	**# of reviews**	**Average # of words/review**	**TripAdvisor median rating**	[1, 1]	[1, 2]	[1, 3]	[1, 4]	[1, 5]	[2, 5]	[3, 5]	[4, 5]	[5, 5]	**Centroid**	**Consensus**
#17	117	82.8	5	0	0	0	0	0	1	8	30	78	{5}	0.895
#18	31	93.5	5	0	0	0	1	0	1	1	3	25	{5}	0.895
#4	36	130.1	5	0	0	0	0	0	0	2	12	22	{5}	0.889
#19	42	135.5	5	0	0	0	0	0	0	3	14	25	{5}	0.881
#12	41	126.3	5	0	0	0	0	0	0	1	18	22	{5}	0.878
#3	87	122.3	5	0	0	0	0	0	0	6	34	47	{5}	0.868
#16	32	76.3	5	0	0	0	0	0	0	4	9	19	{5}	0.867
#7	37	116.2	5	0	0	0	0	1	0	2	13	21	{5}	0.858
#1	35	131.3	5	0	0	0	1	0	0	3	10	21	{5}	0.850
#14	30	113.4	5	0	0	1	0	0	1	1	7	20	{5}	0.850
#6	41	85.7	5	0	0	0	1	0	2	3	10	25	{5}	0.835
#15	30	119.3	5	0	0	0	0	0	2	5	7	16	{5}	0.808
#13	34	128.1	5	0	0	0	1	0	1	6	7	19	{5}	0.801
#2	32	166.8	5	0	0	0	0	0	0	1	16	15	[4,5]	0.875
#5	60	131.3	5	0	0	0	0	0	1	7	22	30	[4,5]	0.838
#20	32	79.1	5	0	0	1	0	1	2	3	11	14	[4,5]	0.773
#10	100	89.2	5	0	0	3	1	1	6	13	31	45	[4,5]	0.770
#8	29	59.3	5	0	0	1	1	0	2	5	7	13	[4,5]	0.733
#9	61	61.4	4	2	0	4	2	0	4	11	17	21	[4,5]	0.664
#11	51	116.2	4	5	2	1	2	0	4	14	10	13	[3,5]	0.559

Table 6: Summary of results

used to differentiate hotels having the same ratings.

The contributions of this paper are threefold. First, it introduces hesitancy in the assessment of each review by means of sentiment analysis. Second the centroid allows us to fuse the information introduced in the reviews and rating. Third, the consensus measure allows us to better understand previous ratings allowing users of recommenders systems to immediately identify which of the hotels will have more variability in their reviews. From a general perspective, the ability to distinguish between items having the same ratings could be beneficial to intelligent personal assistants. Rather than offering a list of the top items based on ratings, an intelligent personal assistant may suggest a single alternative to the user. This scenario would be more reflective of a conversation between friends.

Future research will be focused in two main directions. First, a further study of the properties of the presented consensus degree in comparison with other similar measures will be carried out. Second, some experiments will be run to test the applicability of the methodology in real recommendation scenarios. This second part will consider its applicability in different domains, model performance, and interpretation by real users.

Acknowledgement

This research has been partially supported by the Secretary of Universities and Research of the Department of Enterprise and Knowledge of the Generalitat de Catalunya (2017 DI 086) and by the INVITE Research Project (TIN2016-80049-C2-1-R and TIN2016-80049-C2-2-R (AEI/FEDER, UE)), funded by the Spanish Ministry of Science and Information Technology.

References

[Adomavicius and Tuzhilin, 2005] Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6):734–749, 2005.

[Alonso *et al.*, 2010] Sergio Alonso, Enrique Herrera-Viedma, Francisco Chiclana, and Francisco Herrera. A web based consensus support system for group decision making problems and incomplete preferences. *Information Sciences*, 180(23):4477–4495, 2010.

[Amblee and Bui, 2011] Naveen Amblee and Tung Bui. Harnessing the influence of social proof in online shopping: The effect of electronic word of mouth on sales of digital microproducts. *International Journal of Electronic Commerce*, 16(2):91–114, 2011.

[Chen *et al.*, 2015] Li Chen, Guanliang Chen, and Feng Wang. Recommender systems based on user reviews: the state of the art. *User Modeling and User-Adapted Interaction*, 25(2):99–154, 2015.

[Dichter, 1966] Ernest Dichter. How word-of-mouth advertising works. *Harvard business review*, 44(6):147–160, 1966.

[Dong *et al.*, 2015] Y. Dong, X. Chen, and F. Herrera. Minimizing adjusted simple terms in the consensus reaching process with hesitant linguistic assessments in group decision making. *Information Sciences*, 297:95 – 117, 2015.

[Fahmi *et al.*, 2016] A. Fahmi, C. Kahraman, and Ü. Bilen. Electre i method using hesitant linguistic term sets: An application to supplier selection. *International Journal of Computational Intelligence Systems*, 9(1):153–167, 2016.

[Gavilan *et al.*, 2018] Diana Gavilan, Maria Avello, and Gema Martinez-Navarro. The influence of online ratings and reviews on hotel booking consideration. *Tourism Management*, 66:53–61, 2018.

[Herrera and Herrera-Viedma, 2000] Francisco Herrera and Enrique Herrera-Viedma. Linguistic decision analysis: steps for solving decision problems under linguistic information. *Fuzzy Sets and systems*, 115(1):67–82, 2000.

[M. *et al.*, 2016] Rodríguez R. M., Á. Labella, and L. Martínez. An overview on fuzzy modelling of complex linguistic preferences in decision making. *International Journal of Computational Intelligence Systems*, 9(sup1):81–94, 2016.

[Montes *et al.*, 2015] Rosana Montes, Ana M Sánchez, Pedro Villar, and Francisco Herrera. A web tool to support decision making in the housing market using hesitant fuzzy linguistic term sets. *Applied Soft Computing*, 35:949–957, 2015.

[Montserrat-Adell *et al.*, 2016] J. Montserrat-Adell, N. Agell, M. Sánchez, and F. J. Ruiz. *A Representative in Group Decision by Means of the Extended Set of Hesitant Fuzzy Linguistic Term Sets*, pages 56–67. Springer International Publishing, Cham, 2016.

[Montserrat-Adell *et al.*, 2018] Jordi Montserrat-Adell, Núria Agell, Mónica Sánchez, and Francisco Javier Ruiz. Consensus, dissension and precision in group decision making by means of an algebraic extension of hesitant fuzzy linguistic term sets. *Information Fusion*, 42:1–11, 2018.

[Nielsen, 2011] Finn Årup Nielsen. A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*, 2011.

[Ribeiro *et al.*, 2016] Filipe N Ribeiro, Matheus Araújo, Pollyanna Gonçalves, Marcos André Gonçalves, and Fabrício Benevenuto. Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, 5(1):23, 2016.

[Rodríguez and Martínez, 2015] R. M. Rodríguez and L. Martínez. A consensus model for group decision making with hesitant fuzzy linguistic information. In *2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, pages 540–545, Nov 2015.

[Rodríguez *et al.*, 2012] R. M. Rodríguez, L. Martínez, and F. Herrera. Hesitant fuzzy linguistic terms sets for decision making. *IEEE Transactions on Fuzzy Systems*, 20(1):109–119, 2012.

[Rodríguez *et al.*, 2013] R. M. Rodríguez, L. Martínez, and F. Herrera. A group decision making model dealing with comparative linguistic expressions based on hesitant fuzzy linguistic term sets. *Information Sciences*, 241:28 – 42, 2013.

[Rodríguez *et al.*, 2014] R. M. Rodríguez, H. Liu, and L. Martínez. *A Fuzzy Representation for the Semantics of Hesitant Fuzzy Linguistic Term Sets*, pages 745–757. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[Travé-Massuyès *et al.*, 2005] Louise Travé-Massuyès, Francesc Prats, Mónica Sánchez, and Núria Agell. Relative and absolute order-of-magnitude models unified. *Annals of Mathematics and Artificial Intelligence*, 45(3-4):323–341, 2005.

[Wu and Xu, 2016a] Z. Wu and J. Xu. An interactive consensus reaching model for decision making under hesitation linguistic environment. *Journal of Intelligent and Fuzzy Systems*, 31:1635–1644, 2016.

[Wu and Xu, 2016b] Z. Wu and J. Xu. Possibility distribution-based approach for magdm with hesitant fuzzy linguistic information. *IEEE Transactions on Cybernetics*, 46(3):694–705, March 2016.

[Xiang *et al.*, 2017] Zheng Xiang, Qianzhou Du, Yufeng Ma, and Weiguo Fan. A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58:51–65, 2017.

The company perspective of product variety: a conceptual model

Christina Schabasser
christina.schabasser@live.at

Bert Bredeweg
Informatics Institute, Faculty of Science, University of Amsterdam, The Netherlands
b.bredeweg@uva.nl

Abstract

In economics research questions are often discussed in isolation, thereby forgetting the bigger picture, e.g. regarding product variety, and consumer needs and demands on a product. In this paper, we integrate miscellaneous aspects and establish a conceptual model of changing product variety from a company´s perspective using qualitative reasoning. We use a set of selected publications that gives a contemporary overview of this economic domain and how parts of the system are interconnected. The formalization and visualization of domain knowledge is expected to dissolve discrepancies and misinterpretations in meaning, since it empowers people involved in decision processes to enter in dialog and discussion more easily. The development of causal relationships turned out to be demanding, perhaps also because in the literature there is a lack of answers to some questions. However, that is also a benefit of our research: developing mechanisms that explain how these phenomena may work. The question of how the relevant parts of the system behave if a trigger from the outside comes (or the absence of a trigger) is helpful for thinking about business processes.

Keywords: product variety, customer needs, product range, increasing product variety, conceptual model, causality.

1 Introduction

Choice is good, more choice is better? The answer often depends on the stakeholder who is asking the question. Our conceptual model is about presenting the company's perspective. However, it will not give us an answer to this question either. That is not why we develop this model. The goal of the developed model is to show which causal connections exist, for example, between product variety and costs and sales; knowledge relevant to decision-making processes. Visualization of the causal relationships provides a new perspective and allows to tap into several things at once.

Imagine you are a manager in a company and you know that the market is struggling with changing consumer needs. Now that your competitors are offering a wider range of products to meet your needs, you want to do the same. Your conclusion is a greater choice of products will lead to more customer satisfaction and therefore more sales will be generated.

But would that not be too simplistic? As [Götzfried, 2013] argues, decisions for projects for new products and product variants are company-critical and decisions on product variety are closely tied to almost all other managerial decisions [Gao et al., 2004]. The connections seem to be more nested than assumed. Therefore, decisions on product variety should not be taken lightly and it is worthwhile to spend enough time on it. Decisions made with care consider affected parts as coherent.

Of course, it is also part of our job to think about what parts are outside and within our model. This is important to communicate, so that decision-makers do not get the idea that there are no other influences besides the presented connections in our model. So, the reasons why you as a manager are thinking of a wider range of products can be manifold. The argument that consumers are the ultimate source of demand for product variety [Kim, 2006] is from our point of view too one-sided. Competition among manufacturers [Lee and Schluter, 2002] is mentioned as a further reason for the expansion of the product range. By the way, here you can see that the knowledge is very scattered. Even if consumers are depicted as the only driving force for product diversity in our model, it is important to mention that there are more influences.

We model both the positive (e.g. increasing sales) and negative effects (e.g. increasing costs) of variety. We imagine mechanisms that can bring our model back into balance, and in which the negative sides of product variety outweigh the positive ones. By having a holistic view, we help people get involved in research without having to read all relevant literature.

The different simulation results help to get an idea of how the system could evolve. Our work with the model has shown that certain connections are only imaginable through visualization. In addition, we were occasionally also able to find areas that have not yet been sufficiently investigated. Here, we use assumptions on our part to complete the model. It helps to understand what happens when there are no certain regulatory management mechanisms (such as the technology mechanism).

This paper is structured as follows. Section 2 explains the basic concepts of the used Qualitative Reasoning (QR) software Garp3. Section 3 gives an insight into what is included in our model and in subsections the individual model fragments are presented. Section 4 shows the most important simulation results. Finally, Section 5 summarises the main objectives and reflects on results obtained.

2 Qualitative Modelling with Garp3

Garp3 [Bredeweg et al., 2009] uses entities, agents, assumptions as well as configurations to describe the physical system structure. Quantities, quantity spaces, magnitudes and derivatives, direct influences, proportionalities, correspondences and inequalities are used to describe the system behaviour.

Quantities are the relevant properties of entities that may change under the influence of processes. In contrast to entities, agents are used to model entities outside the modelled system. Configurations represent structural relationships between entities, and entities and agents.

Quantities consist of a quantity value which consists of a magnitude and derivative. The quantity space represents the range of possible values of a quantity. While the magnitude describes the current value of a quantity, the derivative is used to describe its direction of change.

The notion I+ is used to model positive influences [Forbus, 2008], which denote direct relations between two quantities. The notion I- is used for negative influences. Using proportionalities (P+ and P-), the derivative of the target quantities can be determined depending on the derivative of the source quantities. Correspondences (C) are used to model the relations between qualitative values of different quantities. Inequalities (≤, <, 0, >, ≥) are commonly used for indicating that one quantity value is different (or equal) to another quantity value (or dirivative).

Scenarios are applied to model the initial state of a system and serve as input for the qualitative simulator. Model fragments are required to describe the structure of a system and consist of conditions and consequences. Each model fragment represents part of knowledge of the domain that may apply to a certain scenario. The engine searches for model fragments that are applicable to the selected scenario and infers the system behaviour. With the mentioned inputs, different simulation outputs can be generated, including state-graph, value history, equation history and an integrated causal model for each state in the state-graph. States in the state-graph depict qualitatively unique behaviours of the modelled system.

3 Modelling of the problem domain

First of all we discuss the theoretical concepts that form the basis for the content of the presented model, then we move on to modeling and subsequent simulation.

3.1 Conceptual framework

Table 1 gives an overview of the relevant components of the system. It also highlights concepts that are interesting, but not part of our model because they represent a non-mandatory extension or even require a separate model.

According to Table 1, different drivers of product variety can be identified. For the work presented, we decided to focus on the consumer as the driving force. On the one hand, the majority of the articles we analyzed explore relationships between customer needs and product variety of a company and, on the other hand, we agree with Peter Drucker's argument [Webster, 2009] that the customer needs should come first in all situations.

As Table 1 shows, our focus is on the company perspective [Gao et al., 2004] and [Webb, 2011] although the consumer view [Riemenschneider, 2006] and [Kahn, 1998] is to some extent included in the decisions of the company (e.g. company responds to customer needs by adapting its product variety) and therefore in our conceptual model. We understand the company as a closed system and we focus on this system. An extension of the model to the customer perspective would not be expedient. From our point of view, the customer perspective is a separate conceptual model with its own quantities and causal relationships.

3.2 Initial situation

At this point, we model the initial situation of a company facing changing customer needs that are triggered by an external drive. The initial scenario (Figure 1) created in the Garp3 Build environment defines three entities (named Customer, Industry and Management) and one agent (named Society). As explained in Section 2, the agent enables us to model exogenous influences on the system. The entity name was chosen industry (and not just company), because we want to speak to a broad readership of decision makers (for example, individual companies or even participants in a supply chain).

There are three configurations in the scenario (named Member, Engages, and Manages). Configurations are used to define the structural relationships between the entities, and with the agent. Configurations are particularly relevant when searching for model fragments that are applicable to the scenario.

Society has a quantity Drive with magnitude zero. The blue arrows in the quantity spaces show the starting values of their associated quantities. As Figure 1 shows, each qualitative value is either a point (quantity: Drive), or an interval (qantities: Needs, Variety, Costs, Sales, Technology and Production) or not specified in the scenario (quantities: Ratio fit, Ratio profit and Ratio innovation).

Drive is represented as an exogenous quantity (denoted by the exclamation mark in Figure 1). The derivative of drive is influenced by "parabolic positive" (bell-shaped development). Drive may transfer its behavior to other quantities of the system. Customer has been assigned the quantity Needs, with the quantity space Interval, which has only a single value, namely Interval. The accompanying quantities of Industry are: Variety, Costs, Sales, Technology, Production, Ratio innovation, Ratio fit and Ratio profit, whereby the last three have the quantity space {Min, Zero, Plus}.

The derivative quantity spaces (δ) of all quantities in the scenario are unspecified, that is, they are intially unknown.

The scenario starts with needs being equal to variety (shown by the equal sign = between needs and variety).

Costs and sales, technology and production as well as ratio fit and ratio profit have the same value in the initial state. Hence, we start with a balance in the initial scenario (the agent also starts at stable) and therefore there is inititially no reason for a regulatory management actions.

Table 1. Model-relevant research concepts

		References	**Central argument**	**Implementation in the model**
Drivers of product variety	Consumers	[Kim, 2006]	Consumers are the ultimate source of demand for product variety.	Consumers are modelled as an entity with the quantity needs that has a positive influence on the product variety.
	Competition	[Lee and Schluter, 2002]	Separation from competitors through product the number of product variants.	*Not part of the model*
	Information technology	[Brynjolfsson et al., 2010]	There are demand-side technological drivers (e.g. Search and Database Technologies) of changes in product variety.	
Effects of changes in product variety (consumer perspective)	On costs (e.g. more time for decision-making due to high product variety)	[Riemenschneider, 2006]; [Kahn, 1998]	More product variants may entail more time for decision-making in the purchasing proces.	
	On utility	[Riemenschneider, 2006]	More product variants may lead to increased value for the customer.	
Effects of changes in product variety (company perspective)	On information technology	[Gao and Hitt, 2004] Firms seeking to offer greater variety can facilitate this strategy through IT investment.		The quantities product variety, technology and costs are seen in a triangular relationship. A cost-effective situation occurs if technology and product variety are balanced.
	On costs			
	On sales	[Randall and Ulrich, 2001]; [Kim, 2006]	Offering products, which ideally satisfy customer needs, may increase a company's sales.	Sales decrease if customer needs are bigger than product variety.
	On profit	[Webb, 2011]	Product variety creates both problems and opportunities for firms and this affects firm profitability.	There is a mechanism that expresses costs as well as benefits of product variety in the form of the quantity profit.

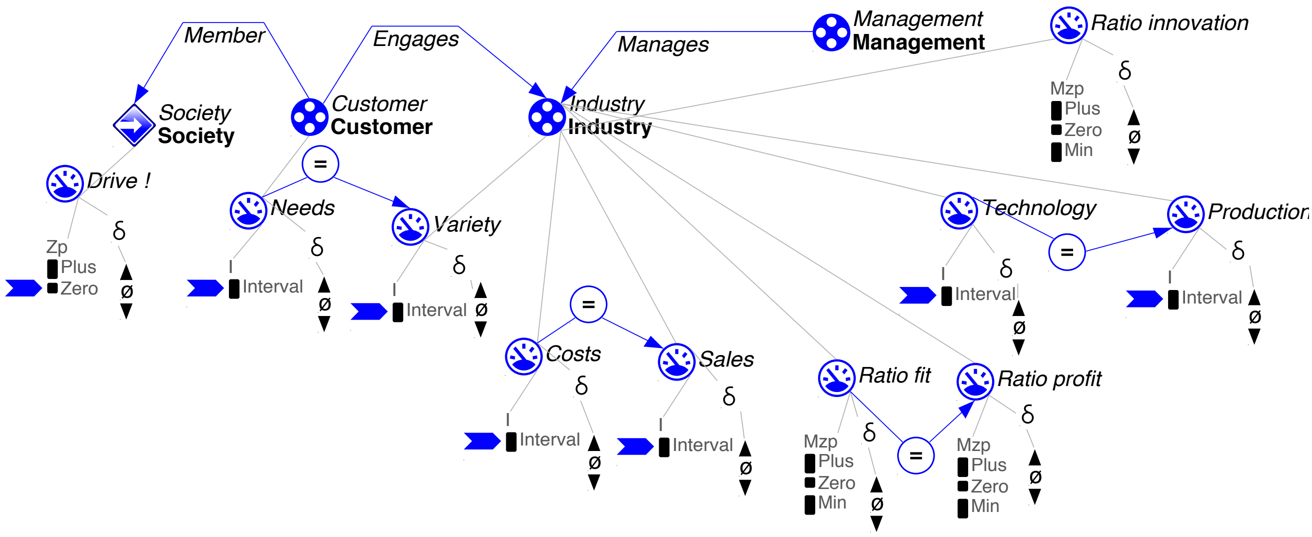


Figure 1. Initial scenario for a company facing changing customer needs.

3.3 Society drive and the effects on critical product variety decisions

We use the exogenous quantity drive which is associated to the agent society (Figure 2). The configuration member from customer to society shows the structural relationship between these two entities. The positive direct influence (I+) between drive and needs is used to express information regarding causality and shows that drive has a positive effect on needs. The I+ causes the quantity needs to increase if the magnitude of drive is positive, decrease if it is negative, and remain steady when it is zero. The derivative quantity spaces (δ) of both the quantities drive and needs are unspecified in the model fragment and change during simulation.

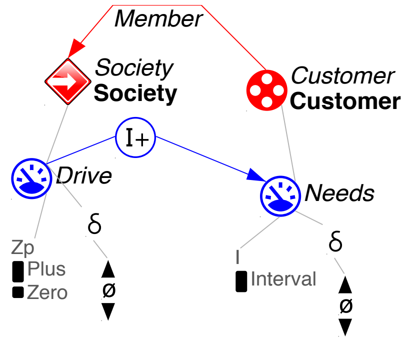


Figure 2. Drive on Needs.

Figure 3 shows a mechanism that tries to balance variety and needs via a management action. The configuration Manages connects the entities Management and Industry. The black part of the model (with the entities industry and customer and the quantities variety, needs and ratio fit) is a static model fragment that has been modeled as such and therefore defines the structure of the system. As we reuse this model fragment within the process model shown in Figure 3 as a condition, we refer to it as an imported model fragment. The imported model fragment is named Variety and needs (coloured red in Figure 3).

There is a calculus specifying: Needs – Variety = Ratio fit. Moreover, Needs has a positive proportionality (P+) with Ratio fit. Therefore changes in needs propagate to changes in ratio fit in the same direction. The negative proportionality (P-) between variety and ratio fit cause ratio fit to decrease if variety increases, ratio fit to remain steady if variety remains steady, and ratio fit to increase if variety decreases.

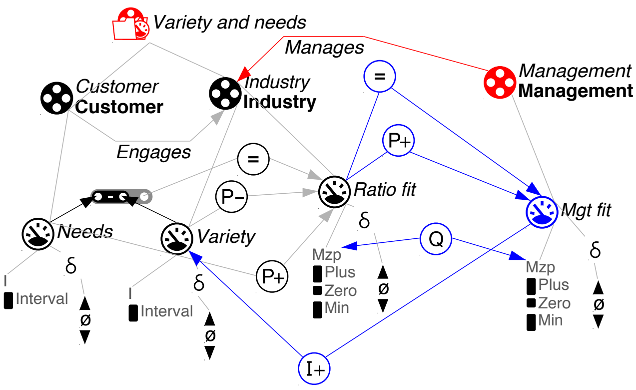


Figure 3. Mechanism management fit.

The entity management with its accompanying quantity Mgt fit has a quantity space {minus, zero, plus} and shows how well the industry meets the needs of its customers with the product range. If needs is bigger than variety, variety has to be increased (denoted by the I+ from Mgt fit to Variety). The fragment specifies a positive proportionality (P+) between Ratio fit and Mgt fit (the latter following changes happening to the former). There is a correspondence (Q) between the quantity spaces of Ratio fit and Mgt fit; specifying that these two quantities have co-occurring magnitudes. The two quantities also have an equality (Ratio fit=Mgt fit).

Figure 4 shows the relationship between Ratio fit and Sales (I- from Ratio fit to Sales). This mechanism triggers sales to fall if customer needs and product variety are not in balance. Therefore, when Ratio fit has value Plus, it makes the Sales decrease, and when Ratio fit has value Zero, it does not change the Sales, and when Ratio fit has value Minus, it makes the Sales increase.

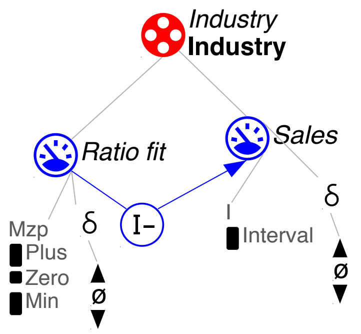


Figure 4. Ratio fit on Sale.

3.4 Product variety and IT as complementaries

Figure 5 shows a mechanism that tries to balance technology and production via a management action. There is a cause-effect dependency (P+) from Production to Ratio innovation, therefore ratio innovation follows changes happening to production. Technology has an indirect negative influence (P-) on Ratio innovation. The negative proportionality will decrease Ratio fit if Technology is increasing, has no effect on Ratio fit if it is stable, and will increase Ratio fit if it is decreasing. The calculus relation (Technology – Production = Ratio innovation) shows the change for Ratio innovation when Technology and Production are out of balance.

If production is bigger than technology, technology has to be increased by management (denoted by the I+ from Mgt innovation to technology). Mgt innovation has quantity space {minus, zero, plus}. Ratio innovation has a positive proportionality (P+) with Mgt innovation. Therefore changes in Ratio innovation propagate to changes in Mgt innovation in the same direction.

The quantity space correspondence (Q) between the quantity spaces of Ratio innovation and Mgt innovation indicates that these two quantities have co-ocurring magnitudes. The two quantities Ratio innovation and Mgt innovation have an equality (Ratio innovation=Mgt innovation).

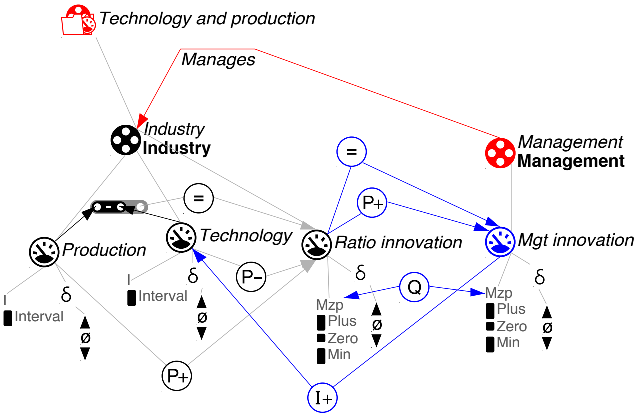


Figure 5. Mechanism innovation management.

Figure 6 shows the causal relationship between Ratio innovation and Costs. Therefore, when Ratio innovation has value Plus, it makes the Costs increase, and when Ratio innovation has value Zero, it does not change the Costs. When Ratio innovation has value Minus, it makes the Costs decrease.

The negative influence (I-) from Ratio fit to Sales will decrease Sales if Ratio fit is Plus, will increase Sales if Ratio fit is Min, and remain stable if Ratio fit is Zero.

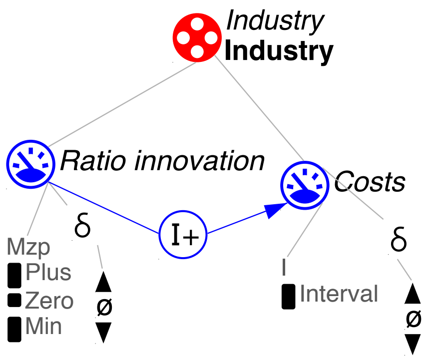


Figure 6. Ratio innovation on Costs

Recall that Figures 3 and 5 show our two management mechanisms. On the one hand Management and the accompanying quantity Mgt fit tries to achieve a balance between Variety and Needs (see Figure 3) and on the other hand Management and the accompanying quantity Mgt innovation tries to balance Technology and Production (as shown in Figure 5).

Variety and needs (use in the model fragment shown in Figure 3) as well as Technology and production (used in the model fragments shown in Figure 5) are imported model fragments that are reused for these mechanisms.

Figures 4 and 6 show the consequence on Sales and Costs, following the balance between Variety and Needs, and Technology and Production. Respectively. Only in the case that the ratios (Ratio fit and Ratio Innovation are zero, there is no change for Sales or Costs.

3.5 Sales and costs balance

Figure 7 depicts the possible effects of Sales and Costs on Ratio Profit. The mathematical calculus (Minus) is used to calculate the difference between Sales and Costs. We define a positive proportionality (P+) from Sales to Ratio profit and a negative one (P-) from Sales to Ratio profit, indicating that a potential increase in Sales would result in an increase in Ratio profit, and an increase in Costs would set the Ratio profit to decrease.

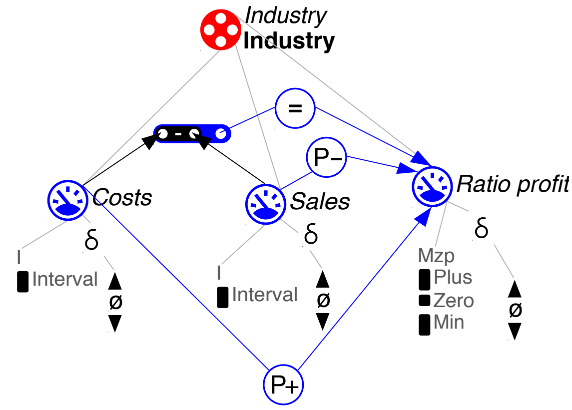


Figure 7. Costs and Sales.

3.6 Simulation results

We use the simulation preference *fastest path* heuristic to avoid overwhelming results. Simulating the initial scenario (Figure 1) produces a state graph with 14 states (see Figure 8) as an end result, whereby state 12 is the only stable end state. Each state reflects a qualitatively distinct behavior of the system. The arrows between two states (e.g. between state 1 and 2) are the state transitions. We select the behaviour path [1 ➛2 ➛ 3 ➛ 4 ➛ 5 ➛ 6 ➛ 12] for further analysis.

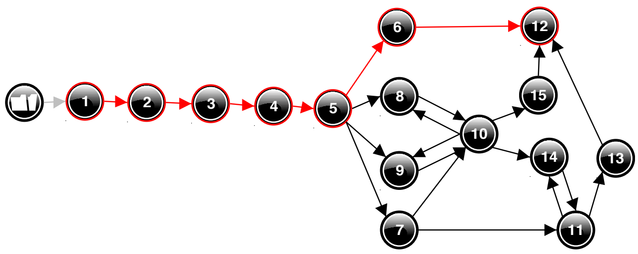


Figure 8. State graph with selected path.

As well as the state graph, the value history provides a particualar view on the simulation results. The value history diagram (Figure 9) enables us to follow the changes each quantity undergoes during the simulation. The selected path as well as all the other paths show the same basic behavior:

Needs increase and Variety tries to follow Needs. As shown in Figure 9, Needs increases from state 2 onwards . Variety follows this rising trend from state 3 on. The diagram presents an increase in Costs (states 4, 5 and 6) as well as a decrease in Sales (states 3, 4, 5 and 6).

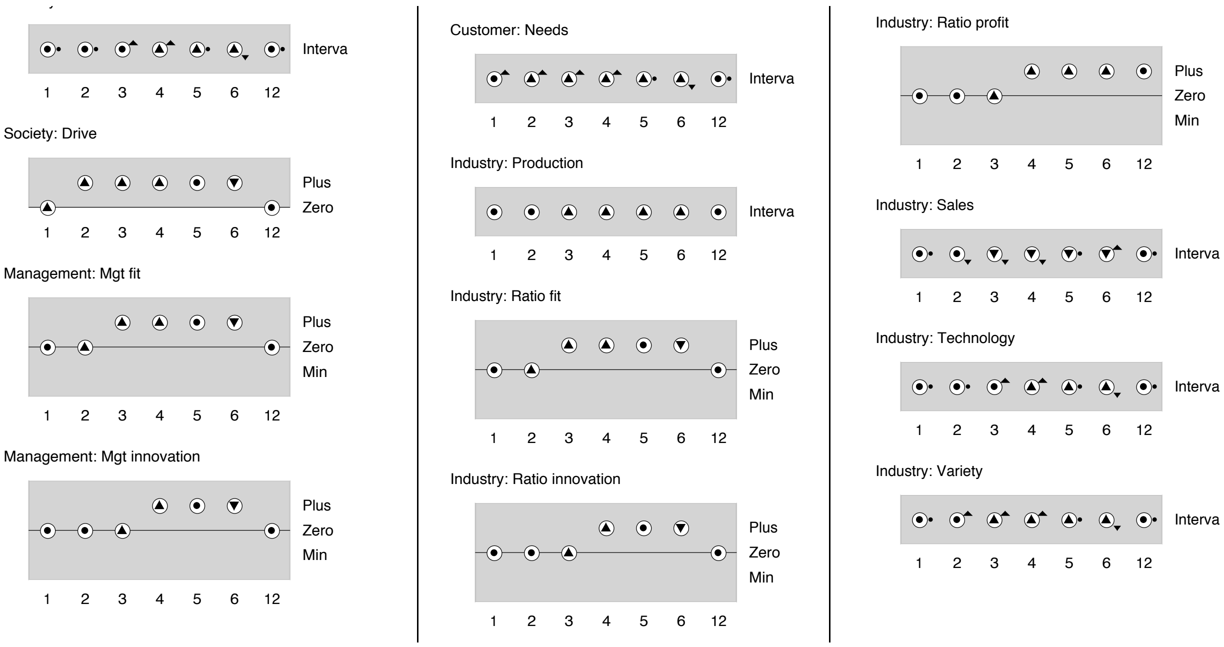


Figure 9. Value history for path [1 → 2 → 3 → 4 → 5 → 6 → 12]

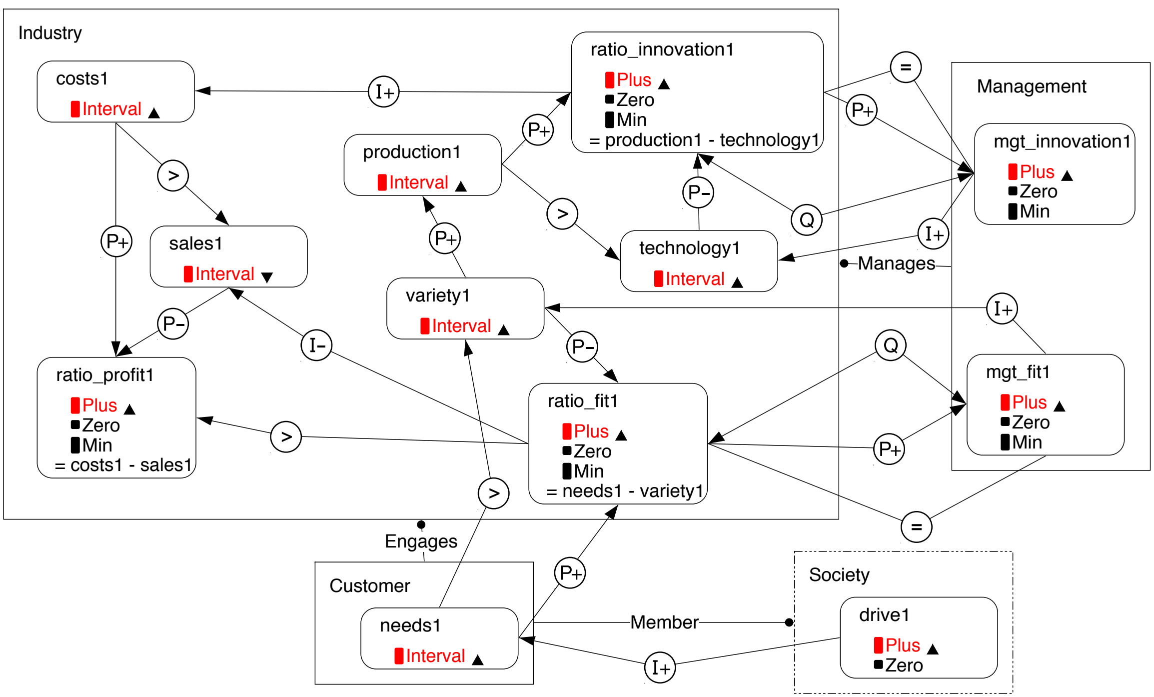


Figure 10. Dependency view for state 4.

The dependency view (also referred to as the causal model) provides an overview of the causal relationships between the quantities in a particlar state (see Figure 10). Once again, it becomes clear at a glance which entities play a role in our conceptual model, namely society, customer, industry and management. The red marking in the individual quantity spaces shows the current value of the quantities in state 4. The black arrows in the quantity spaces show the derivatives. Except for sales, a rising trend is expected for all quantities.

4 Concluding remarks

The main question that initiated this research was how to formalize knowledge about a company's product variety – that is all the products and variants it offers – jand its causal relationships. We started with a literature research to identify key drivers of product variety as well as the interdependencies between product variations and other business sectors. The accurate demarcation of parts inside and outside the system sharpened our common understanding of the domain to be modelled.

The following insights have emerged from the presented work.

- Language is sometimes difficult to understand and leaves room for interpretation. Common concepts are often named differently. Therefore visualization is a good means to overcome this, since it helps to recognize potential errors.
- It was hard to find the right abstraction level, which provides a value added for the user. Distinguishing relevant from irrelevant content turned out to be difficult. For this, one must already have a good overview of causal relationships in order to accomplish this step.

The paper reveals that it is possible to use conceptual modelling for developing ansswers to ecnomical questions. The visualization capabilities of the Qualitative Reasoning (QR) software (Garp3) help to make complex phenomena insightfull and understand them.

For future research, we are working on further management mechanisms. In concrete terms the planned management mechanisms are:

- Marketing management as an awareness creator and therefore the driver of sales.
- Cannibalization management to defend turnover and sales.
- Product bundling management as an effective means to boost profit.

We plan to introduce further quantities in the model, including complexity. In addition, there will be a max value in the quantity space of variety, that should not be exceeded (otherwise sales will be lost).

References

Bredeweg, B., Linnebank, F.E., Bouwer, A.J. and Liem, J. 2009. Garp3 - Workbench for qualitative modelling and simulation. Ecological Informatics, 4(5-6), 263-281.

Brynjolfsson, E., Hu, Y. and Smith, M. (2010). Long Tails versus Superstars. The Effect of IT on Product Variety and Sales Concentration Patterns. In Gupta, Alok (ed.), Information Systems Research 21(4), 736-747. USA: Informs. Retrieved 07.06.2018 from https://doi.org/10.1287/isre.1100.0325

Forbus, K.D. (2008). Qualitative Modeling. In Handbook of Knowledge Representation, Volume 3 of Foundations of Artificial Intelligence, (eds). F. van Harmelen, V. Lifschitz, and B. Porter, 361–393. Amsterdam: Elsevier.

Gao, G. and Hitt, L. (2004). IT and product variety: Evidence from Panel Data. In ICIS 2004 Proceedings (ed.), USA: Twenty-Fifth International Conference on Information Systems. Retrieved 07.06.2018 from http://aisel.aisnet.org/icis2004/

Götzfried, M. (2013). Managing Complexity Induced by Product Variety in Manufacturing Companies. Complexity Evaluation and Integration in Decision-Making. Diss Bamberg: Difo-Druck.

Kahn, B. (1998). Variety: From the Consumer's Perspective In Teck Ho and C. S. Tang (ed.), Product Variety Management: Research Advances, pp.19-37.

Kim, Y.A. (2006). The Impact on Customers' Perception of Product Variety. Korea Review of International Studies, 9 (1).

Lee, C. and Schluter, G. (2002). Why Do Food Manufacturers Introduce New Products? In Journal of Food Distribution Research, 33 (1).

Randall, T. and Ulrich, K. (2001). Product Variety, Supply Chain Structure, and Firm Performance: Analysis of the U.S. Bicycle Industry. In Management Science 47 (12), 1588–1604.

Riemenschneider, M. (2006). Der Wert von Produktvielfalt: Wirkung großer Sortimente auf das Verhalten von Konsumenten. Diss St. Gallen: Deutscher Universitäts-Verlag.

Webb, S. (2011). Product Variety: An investigation into its revenue, cost and profit. Diss Michigan State University

Webster, F. (2009). Marketing IS management: The wisdom of Peter Drucker. In J. of the Acad. Mark. Sci., pp. 20-37.

A Qualitative Gait Model for Diagnosis of Juvenile Idiopathic Arthritis

Peter Struss[1] and Robert Reiser[1] and Florian Kreuzpointner[2]
Technische Universität München, [1]Department of Informatics/
[2]Department of Sport and Health Sciences, Munich, Germany
email: {struss, robert.reiser, kf}@tum.de

Abstract

Juvenile idiopathic arthritis is an autoimmune disorder that causes inflammation and pain in joints and, hence, may lead to posture and movement modifications and muscular imbalance with reduced range of motion in the affected joints. In order to support diagnosis and selection of adequate therapy, certain features of the gait of patients are recorded, based on several markers attached to their legs. In our work, we attempt a model-based solution to the diagnostic interpretation of the recorded gait parameters. The model is a compositional one, with bones, joints, muscles, and the "control" of the muscles (the central nervous system) as the building blocks, and captures deviations of parameters from the nominal range. We present a first version of the model library under certain simplifications and restrictions, esp. the limitation to 2d.

1. Introduction

Component-oriented model-based diagnosis (Struss, 2008) has mainly been designed for and applied to artifacts. One of the main reasons for this is that they (usually) have a clear (designed!) structure and are often constructed using a set of building blocks, which allows for compositional modeling based on a library of reusable models. In contrast, medical diagnosis faces very complex systems, often with no obvious structure, spatially distributed organs and processes etc. In this paper, we describe an attempt to apply qualitative modeling and model-based diagnosis techniques to a medical problem that appears to be closer to the traditional applications.

The context is given by data acquisition and diagnostic techniques in the area of juvenile idiopathic arthritis (JIA), where young patients suffer from inflammation of joints, which may seriously harm their motion capabilities. In particular, their gait may be affected leading to abnormal motions. The commonly used 3d gait analysis records deviations from normal gait as a basis for diagnosis and monitoring of the patients. In order to support this task, we started the development of a compositional (simplified and qualitative) model of human legs and gait, which can be exploited by model-based diagnosis.

Of course, two legs in motion form a complex mechanical system. Inspired by the kind of expert interpretation of the patient data, we start with a qualitative, purely geometrical 2d model of a leg, ignoring the complex kinetics and its control by the central nervous system (CNS). This is what we present in this paper.

We start by describing JIA and its diagnosis and therapy and, in particular, standard 3d gait analysis. Section 3 and 4 introduce the foundations of the model related to the geometrical and mechanical aspects and the control of the motion. The models based on these constraints are presented in section 5. Then we discuss first diagnostic results and open questions and future work.

2. Diagnosis and Therapy of JIA

The EULAR (the European League Against Rheumatism) proposed the term juvenile idiopathic arthritis for the heterogeneous group of disorders that manifest as juvenile arthritis (inflammation of a joint). The definition implies that the arthritis begins before the age of 16 years and lasts for at least 3 months (Sherry et al., 2011).

2.1. The Disease

JIA is the most common chronic arthritis in childhood and adolescence. It is an autoimmune and autoinflammatory disease, i.e. there is an immune response against the constituents of the body's own tissues.

This reaction is characterized by joint inflammation, pain and swelling which have an impact on the muscle function and influences the human gait (Hartmann et al., 2010).

According to the classification of the International League Against Rheumatism (ILAR), there are seven subclasses of JIA (Petty et al., 2004). All have a different impact on the muscle function because they cause restrictions at different levels. As JIA is a very heterogeneous disease, it is very important to know the various factors that influence muscle function.

2.2. Diagnosis and Therapy

The etiology of the disease is not clear yet. Additionally, there is no consistent pattern of joint involvement. This

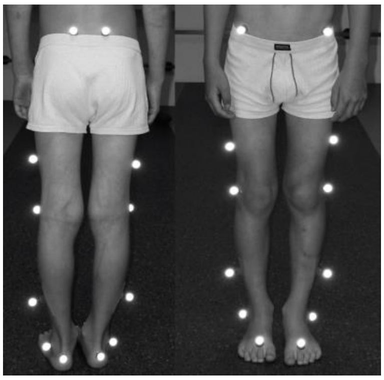


Figure 1 Reflecting markers (Hartmann et al., 2010)

means that under a functional point of view, the disease is very heterogeneous. One subclass, polyarticular JIA (this means that the patient has more than five involved joints), can have six involved joints on the lower extremity or six involved joints on the upper extremity.

The joint involvement has a big impact on the muscle function, which determines the motion sequence of a movement. If the muscle function is restricted there is an anomaly in the motion which limits the patient, for example, in the way they walk.

Patients with JIA are treated within their therapy in a multimodal way. Two major parts of treatment are both medical therapy and functional therapy like physiotherapy and exercise therapy. Drugs treat the inflammatory process, and functional therapy treats the resultant restrictions of the involved joint.

Joint inflammation causes degenerative effects within the joint, which leads to changes of the muscle function.

Pain, dysfunction of the agonist muscles or bony restrictions can affect the ordinary gait function. In terms of treatment planning it is necessary to have a closer look at the outcome of a gait analysis.

2.3. Gait Analysis and its Contribution to Diagnosis and Therapy

Three dimensional gait analysis (3dGA) is a very powerful method to quantify muscle function during walking or running. In general, a standard 3dGA consists of kinematic and kinetic analysis. Some labs use electromyography (EMG), as well. Kinematic parameters are related to joint angular displacements. Kinetic parameters include external ground reaction forces during movement. One can calculate joint loadings with both parameters using inverse dynamics. EMG analyzes the muscular activation that is necessary for quantifying the neuromuscular processes or human locomotion. The most commonly used outcome parameters are kinematic parameters. They are expressed by joint angle displacements.

Today, (3d) gait analysis is a part of a routine procedure in a clinical setting used to quantify movement restrictions to individualize physiotherapy. Basically, gait analysis is performed on an around 10 m long and 3 m wide gait floor in a lab environment. The lab is equipped with a 3d-motion analysis system including at least six infrared cameras, measuring at 120 Hz (e.g. Vicon, MX3) and at least one 3d ground reaction force plate (1080 Hz) (AMTI). The patients are marked in accordance with the Plug-in-Gait Model for the lower extremities (Davis, 1997) with 16 reflecting markers (∅ = 14 mm) (Figure 1).

After post processing the data, the marker trajectories are translated into joint movements during walking, which is normalized to a gait cycle from 0-100%. A gait cycle lasts from heel strike of one side to the next heel strike of the same foot. The main focus lies in the interpretation of the sagittal plane. Figure 2 shows the movement pattern of the ankle joint over one gait cycle in the sagittal plane (i.e. the yz-plane where z denotes the vertical axis and y is the walking direction) which indicates the range of the joint angle during different phases. In this Figure, the vertical axis represents the plantar- and dorsiflexion angle of the ankle joint. The gray-shaded area is the standard deviation around the black solid line, which is the mean of a healthy age matched control group. This is compared to a single patient with JIA. Again the red shaded red area is the standard deviation of the patient with the black solid line which is the mean out of five gait trails. Combining multiple trails of a single patient is done to cover the variability of the gait patterns. In this example, one can see that the patient has a limited maximum plantar flexion, which takes place around push off (vertical line in the Figure). This means that the patient has less potential to accelerate the center of mass during push off.

One of the difficulties of the interpretation is that different causes (ranging across different time scales) may cause the same abnormalities. For instance, if the flexion of a joint is reduced the cause may be

- an inflammation in the joint causes pain for certain extreme ranges of joint angles. As a response, the CNS attempts to confine the range of the joint angle in order to avoid the pain.

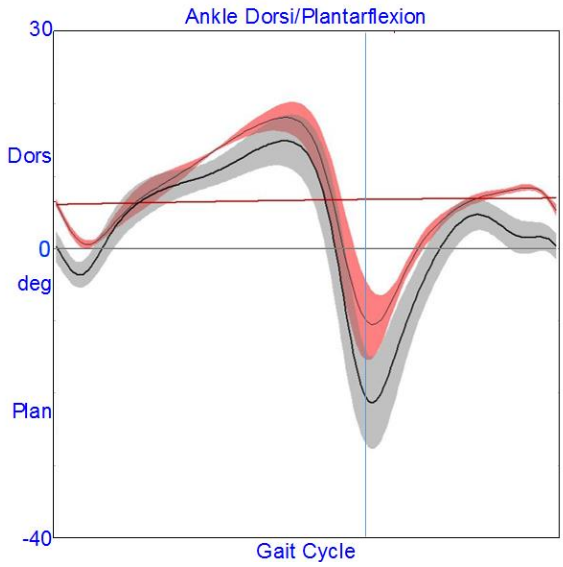


Figure 2 Ankle joint motion in the sagittal plane during walking

- After some time, the modified control regime may actually persist, although the former origin, the inflammation, is no longer present.
- Finally, if the muscle has been restricted in its performance for some time, it may lose its capabilities to operate beyond these restrictions, for instance, achieving a complete contraction.

These causes cannot be discriminated without further information, such as history of the patient's disease, inflammatory values, and additional tests

3. Model-based Gait Analysis – Foundations

It is obvious that computer-based support diagnosis based on the data outlined above has to be knowledge-based, as opposed to data-driven, and the objective of our work is exploring whether and to what extent model-based diagnosis can contribute to the interpretation of the measurements by identifying potential causes for observed deviations of joint study.

In this chapter, the foundations and essential constraints of our model are introduced.

3.1. Component-oriented Modeling

Our approach to modeling the gait is a qualitative, compositional, component-oriented one. This means, we identify different types of entities that have a specific behavior and several instances in the system. These instances ("components") interact with each other, and their combined behaviors establish the overall behavior of the entire system, in a normal or disturbed way.

Developing a compositional model (rather than a black-box model of the entire leg or pair of legs) provides transparency in terms of a clear model structure, supports modifying and extending the model and bears the advantage of re-using model fragments for other purposes (such as differential diagnosis or therapy generation) and other parts of the body.

In our application, the obvious main component types involved in the mechanics of human gait are

- **Bones**
- **Joints**
- **Muscles**

Of course, there are many other parts of the leg involved, but since they are not considered as potential problem sources in our domain, they are neglected. Tendons, for instance, can be treated as part of the muscle or could be seen as the passive connections between bones and muscles, manifested by the respective **terminals** that connect the models of the components and allow sharing of information between them.

In addition to the three main component types, there is the

- **Central nervous system (CNS)**

as the element that determines the gait by controlling the muscles.

3.2. Modeling Decisions and Assumptions

In principle, the human gait is an incredibly complex dynamic process involving, for instance, the interaction of agonist and antagonist muscles (flexor and extensor muscle) in a 3d space. Detailed and numerical modeling has been attempted e.g. in (Koning et al. 2015). Using such models for our purposes appears to be prohibitive due to its complexity. It would also be overly detailed and useless, because the required numerical parameters for tailoring it to an individual patient, and fail to relate to the human diagnostic reasoning. Our model involves several significant simplifications. The expectation that it will serve its purpose is based on the consideration of how human experts perform diagnostic reasoning.

- The **content** of the model: At least a major part of the human diagnostic reasoning seems to ignore the kinetics of the gait. This is why our first model exclusively captures the **geometrical interdependencies only** and ignores forces, momentum, acceleration etc.
- The **granularity** of the model: the description of the gait of an individual patient is based on an abstraction of the numerical measurements to a qualitative level, in determining whether and in which direction angles or positions deviate from envelopes around the nominal curves.
- **Structural reduction**: as already stated above, we obviously have to omit many elements of the body from the model. In particular, we consider **only one pair of muscles for each joint** (muscles come in pairs with one causing the flexion and the other one the extension of the joint when contracting). In reality, there are more muscles affecting the joint. However, in the JIA domain, considering them separately is not feasible and would not contribute to a refined diagnosis, anyway. Another simplification, which is likely to be dropped in the future, is that we **do not consider muscles that work across two joints**.
- **Two dimensions** only: we consider the sagittal plane only, i.e. the one spanned by the vertical axis and the one in the direction of the gait. Also this simplification has to be overcome in more sophisticated models, since avoidance of pain is often achieved by or leads to distortions in the 3rd dimension.

The currently implemented model is additionally limited to

- **one leg** only, which is a restriction, because a deviating motion of one leg may have an impact on the motion of the other one.
- the periods of the gait where it **carries the load** of the body, which is justified by the fact that abnormalities are more likely to show under a load, rather than during the swing phase.

There are other, more or less restrictive modeling assumptions, which we will make explicit when presenting the models in the following.

3.3. The Involved Constraints

3.3.1. Deviation Models

In order to reflect the representation of the gait based on the measurements and its qualitative interpretation as described in section 2, we consider deviations from nominal

coordinates and angles, more specifically, deviations of maximal or minimal angles.
As in (Struss, 2004), for a variable x, we define the **deviation** Δx as the sign of the difference between an actual (or hypothesized) value and a nominal one

(1) $\Delta x := sign\ (x_{act} - x_{nom})$.

For monotonically increasing (decreasing) functions M+ (M-, resp.), we will exploit

(2) $v1 = M^{+}\ (v2) \Rightarrow \Delta v1 = \Delta v2,$
$v1 = M^{-}\ (v2) \Rightarrow \Delta v1 = -\Delta v2$

Based on this, we present the basic constraints that will form the core of the component models, both for the magnitudes of the involved variables (which will not appear in the model, since they are numerical interdependencies) and those for their qualitative deviations that are obtained from them.

3.3.2. Geometric Definitions

In our representation, the vertical axis is z, while x denotes the axis in the direction of the motion (Figure 3). In the representation used, components have terminals as connection points to other components, carrying their shared variables, such as positions. And components have local (state) variables, for instance, the angle of a bone. In our notation, for instance, $Bone.T_{prox}.\Delta x$ refers to the deviation of x in the proximal terminal of Bone, and $Bone.\Delta\alpha$ the deviation of the angle of Bone. Here, the **proximal** terminal (or adjacent component) of a component is the one towards the center of the body, while the other one (towards the foot) is the **distal** one.

- The **bone angle** α is the angle between the vertical axis with the origin placed in its terminal T_{prox} and the bone in counter-clockwise direction (Figure 3).
- The **joint angle** ϵ is the angle between the two connected bones, also in counter-clockwise direction.

3.3.3. Position of a Bone

Based on these definitions, we have for the **magnitudes** of the coordinates of the terminals and the bone angle

(3) $T_{dist}.z = T_{prox}.z + cos(\alpha) * Bone.length$
(4) $T_{dist}.x = T_{prox}.x - sin(\alpha) * Bone.length$

Where Bone.length is the length of the bone.
We introduce landmarks 0, 90, 180 and 270 and the intervals between them as the qualitative domain of α and

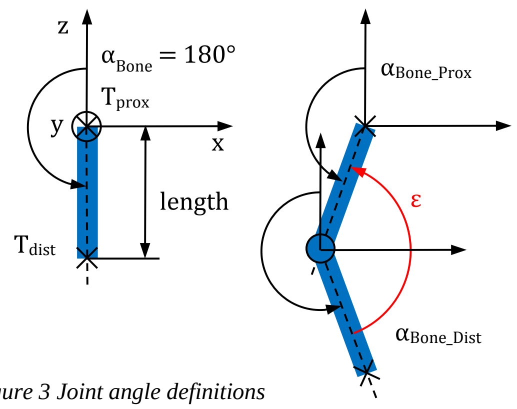


Figure 3 Joint angle definitions

use ⊖ for subtraction in the sign domain. Because cos and sin are piecewise monotonic, if the deviations in α do not cause a crossing of minima or maxima, we obtain the simple constraint between the deviations of α and the coordinates.

(5) $Bone.\alpha_{act} \in \{(0,90), 90, (90,180)\}$
$\wedge\ Bone.\alpha_{nom} \in \{(0,90), 90, (90,180)\}$
$\Rightarrow Bone.\Delta\alpha = Bone.T_{prox}.\Delta z \ominus Bone.T_{dist}.\Delta z$
$etc.$
$\Rightarrow Bone.\Delta\alpha = Bone.T_{dist}.\Delta z \ominus Bone.T_{prox}.\Delta z$

3.3.4. Angles at Joints

According to its definition (see Figure 3), the magnitude of the joint angle ϵ is

(6) $Joint.\epsilon = 180° + Bone_{prox}.\alpha - Bone_{dist}.\alpha$

and its deviation Δϵ

(7) $Joint.\Delta\epsilon = Bone_{prox}.\Delta\alpha \ominus Bone_{dist}.\Delta\alpha$

3.3.5. Impact of Muscle Length

Muscles apply forces to the bones and can modify their relative angles. While we do not explicitly represent the forces in our model, we do represent the monotonic relation between the length of a muscle and the angle difference of the bones it is attached to. If $Muscle_{Front}$ is the one acting on the front side, i.e. in the direction of the motion, and $Muscle_{Back}$ the one in the back, we obtain for the magnitude of ϵ

(8) $Joint.\varepsilon = M^{+}(Muscle_{Front}.length)$
$Joint.\varepsilon = M^{-}(Muscle_{Back}.length)$

Which yields

(9) $Joint.\Delta\varepsilon = Muscle_{Front}.\Delta length$
$Joint.\Delta\varepsilon = -Muscle_{Back}.\Delta length$

for the deviation. Note that this holds independently of which muscle acts as the flexor or extensor, respectively. (This is captured by the possible range of ϵ).

3.3.6. Combination of Muscle Pairs

(8) and (9) indicate that the lengths of the corresponding front and back muscles are coupled; one cannot contract if the other one does not expand. We introduce a variable ∂length to represent the difference to the rest length of a muscle. ∂length=– means contraction (which happens actively), ∂length=+ characterizes an extension (caused by another force, usually the contraction of the antagonist muscle acting on the same joint). Being connected to opposite sides of the same joint, the muscles of a pair are negatively coupled:

(10) $Muscle_{front}.\partial length = -\ Muscle_{back}.\partial length$

And so are the deviations, as captured by (9).
This creates a modeling problem. We cannot simply model the normal behavior of a muscle by stating that it never has a deviating length: if its counteracting muscle suffers from an abnormality and, hence, creates a deviation of the joint angle, the length of the other muscle will be abnormal, too. Thus, it cannot be represented which one is **causing** the deviation. Therefore, we introduce a variable $\Delta length_{pot}$ that represents the **potential** of a muscle regarding its length. If the muscle is working correctly, this deviation regarding the capability of the muscle, will be zero, but its

actual length may deviate, because the potential of the other muscle is abnormal. Only if both potential lengths are normal, the resulting lengths for both will be nominal. This is captured by

$$
\begin{aligned}
(11)\ & Muscle_{Front}.\Delta length = \\
& \quad Muscle_{Front}.\Delta length_{pot} \\
& \qquad IF\ Muscle_{Front}.\Delta length_{pot} \neq 0 \\
& \quad -Muscle_{Back}.\Delta length_{pot} \\
& \qquad IF\ Muscle_{Back}.\Delta length_{pot} \neq 0 \\
& \quad 0 \text{ ELSE}
\end{aligned}
$$

4. Gait Phases

The human gait is a cyclic movement, which is normally represented by a gait cycle (whitch starts with a heel strike and lasts until the upcoming heel strike of the same side). Therefore, a gait cycle consists of a stance phase and a swing phase. Currently, we only consider the stance phase. This phase is further divided into 5 subphases:

(1) Initial contact (I)
(2) Loading response (Load. Resp.)
(3) Midstance (SiSp. Mid)
(4) Terminal stance (SiSp. Term)
(5) Pre-swing

The characterization of these phases in terms of the model variables (angles and coordinates) and their deviations is shown in Table 2.

During the different gait phases, the contribution of the muscles to the respective motion varies. In principle, all muscles are somehow involved in all phases. However, some are crucial to achieving the essential dynamics of a gait phase, while others play a secondary role, e.g. by stabilizing the motion. For instance, during the Midstance phase, an essential motion is stretching the knee, which is achieved by contraction of the – in our model – KneeFrontMuscle, which is the m. quadriceps fermoris. Essential muscle contributions are not confined to contraction. Also, a controlled extension of a muscle may be crucial. While, in a muscle contraction, (neglecting external forces) the fibers themselves can only shorten or remain static, it is also possible that a muscle contracts eccentrically. This happens under an acting external force that is greater than the muscle force when the fibers lengthen while actively creating a resistive force. For example, during the Loading Response Phase the same KneeFrontMuscle, by a resistive extension, prevents a too sudden bending of the knee due to the body weight.

Diagnosis considerations suggest that the muscles that are strongly engaged in achieving the characteristic motion are also the ones that may cause significant modification of the gait when they function or are controlled in a wrong way. In addition, the antagonist muscles of contracting ones may disturb the motions by counteracting in an abnormal way. Hence, we represent each gait phase also by characterizing patterns of muscle activities, which are summarized in Table 3, which lists the essential commands to the different muscles for each gait phase.

5. Component Models

In this section, we introduce the types of terminals that are attached to the component and the domains of the variables used and, based on this, the different component types.

5.1. Domains and Terminal Types

The defined domains are shown in Table 1. CmdDomain captures the different ways the CNS stimulates the muscles. Besides the ones that occur in the normal gait as shown in Table 3, there are commands that aim at limiting the contraction or extension of a muscle in order to limit the resulting joint angle.

PainDomain allows describing whether and in which position pain occurs in a particular joint. Although this is not visible in gait analysis, information about painful motions can be obtained by separate examinations of individual joints.

Table 1 Domains

Domain	Values
PainDom	NoPain, StretchPain, BendPain, BothPain
Sign	-, 0, +
AngleDom	0, (0, 90), 90, (90, 180), 180, (180, 270), 270, (270, 360)
CmdDom	none, Contraction, ResistiveExtension, LimitedContraction, LimitedExtension

There are two types of terminals in the physical system:

- BoneJoint (BJ): Enables sharing of end positions and angle α of bones between them and joints
- JointMuscle (JM): Communicates (deviations of) muscle lengths to the joint and, hence, captures the respective variables

The CNSis connected to joints and muscles through the following terminals:

- Pain (PT) to receive an indication of pain from joints
- Command (CT) to transmit control commands to muscles

The decomposition of the leg model into components of different types (summarized in Table 4) does not reflect exactly the physical objects that constitute the leg. The main reason for this lies in the nature of a joint, which represents the interaction of two bones and (in our simplified version) one muscle pair.

5.2. Joint

A **joint** is considered as a unit comprising the ends of two bones, has two BoneJointTerminals, and determines the relative position and motion of the two connected bones (via constraints (7)). Furthermore, it is connected to one back and one front muscle by two JointMuscleTerminals and mediates their interaction according to constraint (10), although, in reality, the muscles may not be attached close to the joint. The rationale behind this is that we do not want to model forces on the bones and the torque produced (because their endpoints are fixed by the joint).

Table 2 Definition of gait phases

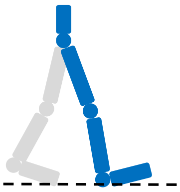
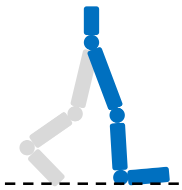
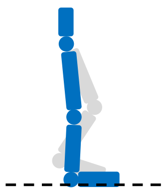
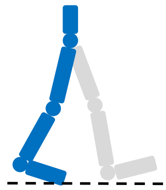
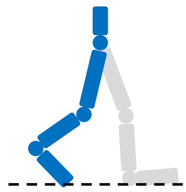
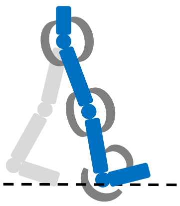
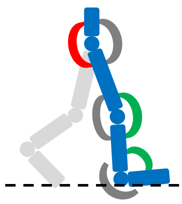
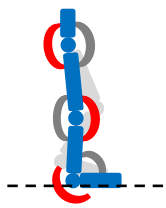
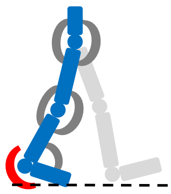
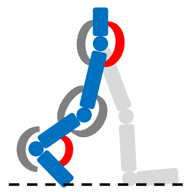
	(1) Init. Contact	(2) L. Response	(3) SiSp. mid	(4) SiSp. term	(5) Pre-swing
					
Pelvis	$\alpha = 180$ $\Delta\alpha = 0$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = 180$ $\Delta\alpha = 0$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = 180$ $\Delta\alpha = 0$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = 180$ $\Delta\alpha = 0$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = 180$ $\Delta\alpha = 0$ $Tprox.z = +$ $Tdist.z = +$
Thigh	$\alpha = (180,270)$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = (180,270)$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = 180$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = (90,180)$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = (90,180)$ $Tprox.z = +$ $Tdist.z = +$
Shank	$\alpha = (180,270)$ $Tprox.z = +$ $Tdist.z = 0$ $Tdist.\Delta z = 0$	$\alpha = (180,270)$ $Tprox.z = +$ $Tdist.z = 0$ $Tdist.\Delta z = 0$	$\alpha = 180°$ $Tprox.z = +$ $Tdist.z = 0$ $Tdist.\Delta z = 0$	$\alpha = (90,180)$ $Tprox.z = +$ $Tdist.z = +$	$\alpha = (90,180)$ $Tprox.z = +$ $Tdist.z = +$
Foot	$\alpha = (270,360)$ $Tprox.z = 0$ $Tprox.\Delta z = 0$ $Tdist.z = +$	$\alpha = (270,360)$ $Tprox.z = 0$ $Tprox.\Delta z = 0$ $Tdist.z = +$	$\alpha = 270$ $Tprox.z = 0$ $Tprox.\Delta z = 0$ $Tdist.z = 0$ $Tdist.\Delta z = 0$	$\alpha = (180,270)$ $Tprox.z = +$ $Tdist.z = 0$ $Tdist.\Delta z = 0$	$\alpha = (180,270)$ $Tprox.z = +$ $Tdist.z = 0$ $Tdist.\Delta z = 0$

Table 3 Discrete muscle activities for the gait phases

	(1) Init. Contact	(2) L. Response	(3) SiSp. mid	(4) SiSp. term	(5) Pre-swing
					
Pelvis stretch	(No phase)	Contraction	Contraction	none	none
Pelvis bend	(No phase)	none	none	none	Contraction
Knee stretch	(No phase)	resistiveExtension	Contraction	none	none
Knee bend	(No phase)	none	none	none	none
Ankle stretch	(No phase)	none	Contraction	Contraction	none
Ankle bend	(No phase)	resistiveExtension	none	none	Contraction

As a consequence, the angle between the bones at the joint is determined by the length of the muscles as indicated by constraint (9), and only one terminal is needed for the connection of each muscle. Also the determination of the muscle length deviations based on the deviations of the potential lengths happens through the joint and, hence, constraint (11) is part of its model (see Figure 4).

Finally, the joint model identifies the coordinate deviations on its two BoneJointTerminals. All these constraints hold unconditionally, i.e. independently of the behavior mode of the joint. These modes are only distinct w.r.t. the existence and type of pain in the joint, which is shared with the central nervous system via the PainTerminal, and accordingly, the “faut modes” are BendPain, StretchPain, BothPain (fixing Tpain.pain to the respective value) and OK (NoPain).

5.3. Bone

A **bone** simply represents the rigid connection of its ends, i.e. the respective joints, carries the resulting constraints (5) (see Table 4) and has no fault models.

5.4. Muscle

A **muscle** is connected to one joint, as depicted in Figure 4, and receives commands from the CNS via its CommandTerminal. The constraints for its different behavior modes have to capture the response of the muscle to these different commands which is then shared with the connected joint. For instance, in its nominal model, we have the tuple

(Tcmd.command=Contraction, ∂length=-, $\Delta length_{pot}$=0),

i.e. the muscle contracts as requested, while the fault mode LimitedContraction contains

(Tcmd.command=Contraction, ∂length=-, $\Delta length_{pot}$=+),

i.e. it does not fully contract. They share the tuple

(Tcmd.command=LimitedContraction, ∂length=-, $\Delta length_{pot}$=+).

While this deviation only shows when the muscle is actively commanded, the failure mode LimitedExtension may have an impact also when it is passive (i.e. receives command none): when its antagonist is contracting, it will restrict its effect due to its limited length. The model contains

$\partial length=+ \Rightarrow \Delta length_{pot}=-$,

which holds under the command ResistiveExtension, but also for Command = none, if the antagonist muscle contracts and causes an extension.

5.5. Central Nervous System

The CNS controls the gait by commanding the muscles appropriately in the various gait phases. It issues commands to the muscles via six CommandTerminals and receives pain signals from the joints via three PainTerminals. Its model is an association of muscle activation patterns with gait phase potentially influenced by indications of pain. Table 3 can be read as the proper association if there is no pain signaled. Otherwise, dependent on which joint indicates which kind of pain, the commands are modified, primarily to their limited version. For instance, in our example, if there is BendPain in the ankle, in gait phase 5, the command to the ankle.FrontMuscle, i.e. the flexor muscle can be modified to LimitedContraction (cf. Table 3).

Table 4 Components

	Joint	Bone	Muscle
Terminals	Tprox (BJ), Tdist (BJ), Tfront (JM), Tback (JM), Tpain (PT)	Tprox (BJ), Tdist (BJ)	Tjoint (JM), Tcmd (CT)
State Vars.	$\Delta\epsilon$ (Sign)		
Constraints	(7), (9), (10), (11)	(5)	

Currently, the only fault model of the CNS is unrestricted behavior, which covers any inadequate control, e.g. due to previously necessary, but meanwhile obsolete compensatory actions. Incorporating this kind of information would require exploiting the anamnesis of the individual patients. CNS are connected to the respecitve muscles (see Figure 4).

6. Diagnosis with the Model – First Results

The model library described above has been implemented in Raz’r (OCC'M Software GmbH, 2011). The overall system model is established as a sequence of bones and joints from the hip to the foot with muscles connected to the back and front of each joint. In addition, the PainTerminal of each joint is connected to the CNS to transmit pain experienced, and the six CmdTerminals of the CNS are connected to the respective muscles.

Furthermore, we introduced a virtual component, “Examination”, that serves as a container to add observations obtained from specific examinations and lab data. Technically, it is connected to the terminals of the other components, esp. the PainTerminals. Currently, it captures inflammatory values (if they are not abnormal, this implies the absence of inflammations and, hence, pain in the joints) and the results of tests of individual joints that would reveal pain in specific positions. In future extensions, this may also include explicit information about individual muscles. All this is meant to help confining the diagnoses obtained from the interpretation of the gait analysis results alone.We evaluated the model using Raz’r’s consistency-based diagnosis engine based on several scenarios that correspond to qualitative deviations of joint angles (in case they lie outside the normal range, as illustrated in Figure 2) and positions that can be extracted manually or, in future solutions, automatically from 3dGA data of patients. This is done for the various gait phases described in section 4. Additional observations can be added via the Examination component explained above. Such cases were

- Gait Phase 4, Ankle.$\Delta\epsilon$ = –
- Gait Phase 2, Knee.$\Delta\epsilon$ = –
- Gait Phase 4, Hip.$\Delta\epsilon$ = +
- Gait Phase 2, Ankle.$\Delta\epsilon$ = +

Despite the simple structure of the model and the involved modeling assumptions, the results obtained so far suggest that the purely geometrical model is able to reproduce an

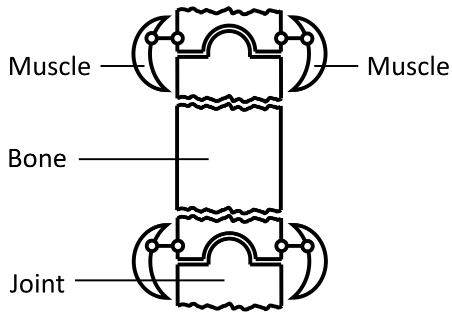


Figure 4 Components

essential part of the conclusions drawn from the data by human experts. The model is able to generate plausible diagnoses, usually expectedly alternative ones. For instance, in the second of the above cases, model-based diagnosis hypothesizes that the muscle KneeFront suffers from LimitedExtension or WrongControl of CNS (excluding a response to pain in the knee). This delivers the kind of information needed to determine additional examinations (in the documented case, the doctor's conclusion was "Function check of both knee muscles").

The introductory case from section 2 (reduced angle at the ankle in phase 4) illustrates limitations of the current solution: the conclusion about the reduction of the push-off force and, hence, of the step width is beyond the scope of the current model: it does not capture forces and momentum, does not handle the swing phase, and would only be able to yield deviations within a single gait phase.

After consolidating the model, we will start an evaluation based on a larger set of cases that have been recorded together with the results of their interpretation by human experts at the German Center for Pediatric and Adolescent Rheumatology, Garmisch-Partenkirchen.

7. Discussion and Future Work

As mentioned above, the model is based on a number of modeling decisions and simplifying assumptions. Some of them appear appropriate for gait analysis and not restrictive. Others are, and we plan to drop them in order to overcome limitations. This will not follow academic motivations related to completeness or theoretical correctness, but be driven by necessities for obtaining better und more useful diagnoses. The planned evaluation will yield criteria for prioritizing modifications and extensions. Some candidates for restrictions to be overcome are, in the order of our current priorities,

- **Restriction** to 2d: sometimes, the impact of an abnormality or the attempt to avoid resulting pain can only be described and observed in the dimension we ignored in our model (e.g. bending a foot inward). Even more fundamentally, the 2d description of the motion of the hip seems quite inadequate.
- **Independent analysis of the two legs**: it seems quite obvious that an abnormality of one leg can affect the other one. For instance, if a leg in the Midstance phase is not fully stretched the swinging of the other leg may have to be modified.
- **Muscles across one joint only**: there are relevant contributions by muscles that act on non-adjacent bones, i.e. have an impact on two joints.
- **Independent analysis of gait phases of one leg**: a deviation created in a certain gait phase may carry over to or affect the following one. It may even be the case that certain characteristic parts of the motion of a gait phase may shift in time and into the next gait phase.
- **Restriction to the load phases**: we have not yet applied the model to the swing phases of the leg, because they usually do not exhibit significant abnormalities (in our context). Actually, this may challenge the model, because it lacks a notion of forces and momentum. Actually, the fact that the foot is fixed during the load phases adds a strong constraint to the model and the observations.

Acknowledgements

The authors would like to thank Josephine Merker, Matthias Hartmann (Motion Analysis Lab) and Prof. Dr. J.-P. Haas (Medical Director) from the German Center for Pediatric and Adolescent Rheumatology for providing one example from a patient (Figure 2) and the discussions. Also special thanks to Florian Grigoleit for his input.

References

(Davis, 1997) Davis, R.B., 1997. Reflections on clinical gait analysis [online]. *Journal of Electromyography and Kinesiology*, **7**(4), 251-257. Available from: 10.1016/S1050-6411(97)00008-4

(Hartmann et al., 2010) Hartmann, M., F. Kreuzpointner, R. Haefner, H. Michels, A. Schwirtz, and J.P. Haas, 2010. Effects of juvenile idiopathic arthritis on kinematics and kinetics of the lower extremities call for consequences in physical activities recommendations. *International journal of pediatrics*, **2010**.

(Koning et al. 2015) Koning, B.H.W., Krogt, Marjolein M. van der, C.T.M. Baten, and B.F.J.M. Koopman, 2015. Driving a musculoskeletal model with inertial and magnetic measurement units [online]. *Computer Methods in Biomechanics and Biomedical Engineering*, **18**(9), 1003-1013. Available from: 10.1080/10255842.2013.867481

(OCC'M Software GmbH, 2011) OCC'M Software GmbH, 2011. FMEA, Testing, Monitoring and Diagnosis (on-board and off-board) are typical applications [online]. Raz'r Model Editor Version 3.0.1107.0 [viewed 2 February 2018]. Available from: http://www.occm.de/

(Petty et al., 2004) Petty, R.E., T.R. Southwood, P. Manners, J. Baum, D.N. Glass, J. Goldenberg, X. He, J. Maldonado-Cocco, J. Orozco-Alcala, A.-M. Prieur, M.E. Suarez-Almazor, and P. Woo, 2004. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *The Journal of rheumatology*, **31**(2), 390-392.

(Sherry et al., 2011) Sherry, D.D., A.R.S. Bhaskar, M. Poduval, C.E. Rabinovich, F. Talavera, M.S. Kocher, L.K. Jung, and B.L. Myones, 2011. *Juvenile Idiopathic Arthritis* [online] [viewed 2 February 2018]. Available from: https://emedicine.medscape.com/article/1007276-overview#aw2aab6b2b4aa

(Struss, 2008) Struss, P., 2008. Model-based Problem Solving. In: F. von Harmelen, V. Lifschitz, and B. Porter, eds. *Handbook of knowledge representation.* Amsterdam: Elsevier, pp. 395-465.

(Struss, 2004) Struss, P., 2004. Models of Behavior Deviations in Model-based Systems. In: R.L.d. Mántaras, ed. *Proceedings / ECAI 2004, 16th European Conference on Artificial Intelligence. August 22 - 27, 2004, Valencia, Spain.* Amsterdam: IOS Press, pp. 883-887.

Non-constructive interval simulation of dynamic systems

Wei Pang, Allan M. Bruce, George M. Coghill
Department of Computing Science
University of Aberdeen, UK
{pang.wei, g.coghill}@abdn.ac.uk

Abstract

Inspired by non-constructive simulation developed in qualitative reasoning, we present a non-constructive interval simulation algorithm for simulating dynamic systems. We recast two integration methods from traditional numerical simulation to make them suitable for non-constructive interval simulation. We then proposed an iterative interval narrowing algorithm to control the growth of intervals during simulation, and we also designed several simulation modes. The simulation algorithm has been both theoretically and experimentally validated.

1 Introduction

Numerical simulation of dynamic systems has been widely used in many engineering and scientific fields [Angermann, 2011]. It numerically simulates differential equation models, such as Ordinary Differential Equation (ODE) and Differential Algebraic Equation (DAE) models, thereby describing how the dynamic system evolves with time.

However, for many real-world problems, due to incomplete knowledge and data, the initial values of model variables and/or the model parameter values have to been given as intervals with lower and upper bounds. Simulating differential equation models as such necessitates the use of *interval analysis* [Moore, 1966], which led to the development of *interval simulation*. In history Interval simulation has attracted the interest of both Qualitative Reasoning (QR) [Kuipers, 1994] and Numerical Simulation (NS) communities. This resulted in the parallel development of interval simulation (in QR the subfield is called semi-quantitative simulation), and therefore there is a gap between these two fields.

In this research we aim to bridge this gap through a novel interval simulation framework. This framework is based on QR but employs techniques developed in NS. Furthermore, we have designed the simulation algorithm to be *non-constructive*, a simulation approach originating from QR (see details in Section 3). This enables the proposed approach to straightforwardly handle models with algebraic loops, which are normally dealt with by NS algorithms through additional, and often complicated and unreliable, operations [Cellier, 1991a].

In the rest of the paper, we first describe the development of interval simulation within the NS and QR fields in Section 2. Then in Section 3 we present the motivations of the research. In Section 4 we introduce the *Morven* formalism used in our approach. In Section 5 the proposed non-constructive interval simulation approach is described in detail. In Section 6 we provide the theoretical analysis on the proposed simulation algorithm. This is followed by the report of a series of experiments in Section 7. Finally Section 8 concludes the paper and explore the future work.

2 Background

2.1 From Numerical Simulation to Interval Simulation

Researchers in NS extended IVP (Initial Value Problem) for ODEs to the interval IVP for ODEs as follows:

$$Y'(t) = F(Y), \tag{1}$$
$$Y(t_0) = Y_0. \tag{2}$$

In the above, $Y \in \mathbb{IR}^n$ is an unknown n-dimensional interval-valued vector variable, where $\mathbb{IR}$ denotes the set of real intervals. $Y_0 \in \mathbb{IR}^n$ is the given initial values. $Y'(t)$ is the first derivative of Y with respect to time t, and $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a given function.

Some researchers further consider situations when the parameters are also intervals, which lead to the replacement of Equation (1) by the following equation:

$$Y'(t) = \mathcal{F}(Y, \theta). \tag{3}$$

In the above, $\mathcal{F}$ is a vector function containing interval-valued parameters, and $\theta \in \mathbb{IR}^m$ is the parameter vector with m being the number of parameters in the model.

Many efforts in NS have been made to recast numerical ODE solvers to deal with interval IVPs for ODEs described by Equations (1) $\sim$ (2) [Nedialkov *et al.*, 1999].

Apart from ODEs, DAEs (Differential Algebraic Equation) are often used in practice. A set of DAEs is represented as follows:

$$F(t, y, y', y'', ..., y^{[m]}) = 0 \tag{4}$$

In the above y is an unknown n-dimensional vector variable, and y', y'',..., $y^{[m]}$ are derivatives of y with respect to time t. Function F is a mapping $R^{n \cdot m+1} \rightarrow R^n$.

The initial values at time t_0 is a solution of DAEs in the

Table 1: Some qualitative constraints in *Morven* and their corresponding mathematical equations

Morven Constraints	Mathematical Equations
sub (dt 0 Z, dt 0 X, dt 0 Y)	$Z(t) = X(t) - Y(t)$
mul (dt 0 Z, dt 0 X, dt 0 Y)	$Z(t) = Y(t) * X(t)$
div (dt 0 Z, dt 0 X, dt 0 Y)	$Z(t) = X(t)/Y(t)$
func (dt 0 Y, dt 0 X)	$Y(t) = f(X(t))$
sub (dt 1 Z, dt 0 X, dt 0 Y)	$dZ(t)/dt = X(t) - Y(t)$
func (dt 1 Y, dt 0 X)	$dY(t)/dt = f(X(t))$

Table 2: The *Morven* Model for the Single Tank System

Differential Plane 0	
C1: mul (dt 0 q_o, dt 0 k, dt 0 V)	$(q_o = kV)$
C2: sub (dt 1 V, dt 0 q_i, dt 0 q_o)	$(V' = q_i - q_o)$
Differential Plane 1	
C3: mul (dt 1 q_o, dt 1 k, dt1 V)	$(q_o' = kV')$
C4: sub (dt 2 V, dt 1 q_i, dt1 q_o)	$(V'' = q_i' - q_o')$

form of Equation (4) in the following form:

$$F(t_0, y(t_0), y'(t_0), y''(t_0), ..., y^{[m]}(t_0)) = 0 \qquad (5)$$

The IVP for DAEs is to numerically simulate DAEs in the form of Equation (4) given the initial values in Equation (5). Similar to the way we represent the interval IVPs for ODEs, the interval IVPs for DAEs can be formulated according to Equations (4) and (5). The development of interval DAE solvers also received some attention [Nedialkov and Pryce, 2005] recently.

2.2 Semi-quantitative Simulation

In the Qualitative Reasoning (QR) community, qualitative simulation uses Qualitative Differential Equations (QDEs) [Kuipers, 1994; Coghill, 1996; Bruce and Coghill, 2005]. Table 1 lists some *Morven* constraints and the corresponding mathematical equations.

Some researchers started from qualitative simulation engines such as QSIM and incorporated incomplete quantitative information (in the form of intervals), and this led to the development of *semi-quantitative simulation* algorithms, for instance, Q2 [Kuipers and Berleant, 1988] and Q3 [Berleant and Kuipers, 1997].

3 Motivations of the Research

From Section 2 we can see that the researchers from the NS and QR communities took different approaches to interval simulation, and there is a gap between these two communities.

From the QR perspective, there is room for improvement for interval simulators: first, as far as the authors are aware, up till now only the Euler and Runge-Kutta integration methods have been used in QR. So one motivation of our research is to explore the potential of a wider range of integration techniques as applied to QR-based interval simulation.

Second, there are both constructive and non-constructive simulation [Wiegand, 1991; Coghill and Chantler, 1999] in QR. Constructive simulation requires the explicit forms of the derivatives of variables. In constructive simulation, at time point t_i, the derivatives are first calculated directly from their explicit forms, and then used for the estimation of the magnitudes of variables at the next time point t_{i+1}. Constructive simulation has been extensively used in NS, but one limitation is that when models are given as implicit forms (i.e., DAEs), additional procedures have to be performed to obtain the explicit forms, and such procedures may be computationally expensive. In addition, to solve a DAE constructively, the structure analysis may fail in some ill-structured DAEs [Pryce, 2001], which makes it impossible to obtain the explicit forms. Furthermore, it is not straightforward for constructive simulation to deal with models containing algebraic loops, which are sometimes used when modelling some real-world systems [Cellier, 1991b].

Within QR there are a number of non-constructive qualitative simulators, such as QSIM and FuSim [Shen and Leitch, 1993]. Non-constructive simulation essentially employs a generate-and-eliminate strategy, and it does not require the explicit forms of variables or their derivatives. This is particularly effective when dealing with some ill-structured DAEs or DAEs containing algebraic loops.

In interval simulation, variables take interval values instead of qualitative values. However, under the QR framework a differential equation model with interval initial conditions and parameter values can be converted into a semi-quantitative model composed of several constraints. So similar to qualitative simulation, in interval simulation we can use each constraint to determine the ranges of interval values for all variables of this constraint (generate). Then the resulting interval value for a variable will be the intersection of all intervals obtained from each individual constraint (eliminate). This means we can also perform the interval simulation in a non-constructive manner.

The above consideration leads to another motivation of this research: implement a non-constructive interval simulator. We expect such research will contribute to both the NS and QR communities.

4 The *Morven* Framework

In this research we use the *Morven* [Coghill, 1996] formalism to represent semi-quantitative models. The *Morven* framework is a constraint-based fuzzy qualitative system. Qualitative constraints in a *Morven* model are distributed over multiple *differential planes*. Qualitative variables in *Morven* are in the form of variable length vectors.

Consider the ODE model of the single tank system:

$$q_o = kV, \qquad (6)$$

$$dV/dt = q_i - q_o, \qquad (7)$$

where V is the volume of the liquid in the tank, q_i is the inflow, q_o is the outflow, and k is a positive constant coefficient.

The above ODE model can be converted into a *Morven* model shown in Table 2. It is noted that *Morven* can use function constraints (”func”) to represent qualitative models.

5 Non-constructive Interval Simulation

There are two phases in non-constructive qualitative simulation [Coghill and Chantler, 1999]: Transition Analysis (TA) and Qualitative Analysis (QA), which correspond to interval integration and interval refinement in interval simulation, respectively, and they are two most important components of non-constructive interval simulation.

5.1 Interval Arithmetic

The interval arithmetic used in our algorithm is defined in Table 3.

Table 3: Interval Arithmetic Operations

Let:	$m = [a \quad b], n = [c \quad d]$	
Operation	Result	Conditions
$-n$	$[d \quad c]$	all n
$\frac{1}{n}$	$[\frac{1}{d} \quad \frac{1}{c}]$	$c, d > 0$ or $c, d < 0$
	$[-\infty \quad \infty]$	$c \leq 0$ and $d \geq 0$
$m + n$	$[a + c \quad b + d]$	all m, n
$m - n$	$[a - d \quad b - c]$	$m \neq n$
	$[0 \quad 0]$	$m = n$
$m \times n$	$[ac \quad bd]$	$m = n$ and $(c, d > 0$ or $c, d < 0)$
	$[0 \quad bd]$	$m = n$ and $c \leq 0$ and $d \geq 0$
	$[ac \quad bd]$	$m \neq n$ and $a, c > 0$
	$[bc \quad ad]$	$m \neq n$ and $a > 0$ and $d < 0$
	$[bc \quad bd]$	$m \neq n$ and $a > 0$ and $c \leq 0$ and $d \geq 0$
	$[ad \quad bc]$	$m \neq n$ and $b < 0$ and $c > 0$
	$[bd \quad ac]$	$m \neq n$ and $b < 0$ and $d < 0$
	$[ad \quad ac]$	$m \neq n$ and $b < 0$ and $c \leq 0$ and $d \geq 0$
	$[ad \quad bd]$	$m \neq n$ and $a \leq 0$ and $b \geq 0$ and $c > 0$
	$[bc \quad ac]$	$m \neq n$ and $a \leq 0$ and $b \geq 0$ and $d < 0$
	$[min(ad, bc) \quad max(ac, bd)]$	$m \neq n$ and $a \leq 0$ and $b \geq 0$ and $c \leq 0$ and $d \geq 0$
$\frac{m}{n}$	$[1 \quad 1]$	$m = n$
	$m \times \frac{1}{n}$	$m \neq n$
$m \cap n$	$[max(a, c), min(b, d)]$	$a \leq c \leq b$ or $c \leq a \leq d$
	$\emptyset$	$b > c$ or $d > a$
$m \cup n$	[max(a, c), min(b, d)]	

$m \neq n$ denotes that the intervals do not correspond to the same interval whereas $m = n$ indicates that the intervals do correspond to the same interval. $a, c > 0$ indicates that both intervals are positive whereas $b, d < 0$ dictates that both intervals are negative. $c \leq 0$ and $d \geq 0$ (as well as $a \leq 0$ and $b \geq 0$) governs that the interval spans zero. It is possible to define $m \times n$ for when both intervals span zero however it has been left out in this table for simplicity.

5.2 Integration Methods

In this subsection, we investigate which integration methods can be used in non-constructive simulation as $Morven$ models can represent those ill-structured ODE and DAE models which are impossible to simulate constructively.

We have explored several common integration methods: (1) Euler methods [Ascher and Petzold, 1998] (including both forward and backward Euler methods), (2) Runge-Kutta methods [Butcher, 2008], (3) Taylor Series Expansion [Arfken *et al.*, 2005], (4) the linear multi-step methods [Butcher, 2003a], including the Adams-Bashforth (AB) methods [Butcher, 2003b], Adams-Moulton and Backward Differentiation Formulas (BDF) [Hairer *et al.*, 1993], (5) the predictor-corrector methods [Press *et al.*, 1992], including the Euler Trapezoidal method and the Adams-Bathforth-Moulton method [Mathews and Fink, 2004].

Among all the above mentioned methods, we identified that Taylor Series (with the forward Euler method being a special case) and AB methods can be used in non-constructive simulation. They can also be easily recast for interval simulation by using the interval mathematics defined in Section 5.1. Therefore, in this research these two integration approaches will be investigated.

We take the Euler methods as an example to demonstrate how we determine the suitability of an integration method for non-constructive simulation, and due to page limit, we will not present the investigation on other integration methods.

We study two streams of Euler methods: forward and backward. The forward Euler method for constructive simulation is given as follows:

$$y_{n+1} = y_n + hf(y_n) \tag{8}$$

In the above y_n is the magnitude of y at time step t_n; $f(y_n)$ is the explicit form of y'_n given by the model: $y'_n = f(y_n)$; y_{n+1} is the magnitude of y at the next time step t_{n+1}, and h is the step size $(t_{n+1} - t_n)$. In constructive simulation, at time step t_{n+1}, first we calculate the value of y_{n+1} according to Equation (8), then this value will be used to calculate the value of y'_{n+1} in the next simulation step t_{n+1} by evaluating $f(y_{n+1})$.

In the context of non-constructive simulation, at time step t_{n+1}, the value of $f(y_n)$ cannot be calculated directly as the form of $f(y_n)$ is not explicitly given. However, the value of y'_n can be taken from previous calculation at time step t_n, and replace $f(y_n)$ in Equation (8), which is shown below:

$$y_{n+1} = y_n + hy'_n \tag{9}$$

In addition, the initial value of y'_0 is either given or calculated by the interval narrowing algorithm which will be described in Section 5.3. This means that the forward Euler method can be used in non-constructive simulation.

The backward Euler method for constructive simulation is given as follows:

$$y_{n+1} = y_n + hf(y_{n+1}), \tag{10}$$

where $f(y_{n+1})$ is the explicit form of y'_{n+1}: $y'_{n+1} = f(y_{n+1})$. In constructive simulation the above equation has to be solved to obtain the precise value of y_{n+1}, for instance, by the fixed point iteration method [Burden and Faires, 2000].

On the other hand, in non-constructive simulation at the time step t_{n+1} the explicit form $f(y_{n+1})$ cannot be obtained straightforwardly, and the value of y'_{n+1} is not calculated yet. This means the backward Euler is not suitable for non-constructive simulation.

5.3 Interval Narrowing Algorithm

Having defined the interval mathematics and chosen the integration methods, the next component to be developed is the interval narrowing algorithm. The interval narrowing algorithm is to control the growth of intervals during simulation, which is achieved by iteratively applying the model constraints to intervals.

We first propose the Inverse Constraint Operations, as detailed below: for each constraint in the model, we obtain all of its mathematically equivalent forms, each of which is called an inverse of this constraint in this paper. For example, consider the following constraint:

$$A = B + C, \tag{11}$$

its inverses will be the following two:

$$B = A - C, \tag{12}$$

$$C = A - B. \tag{13}$$

Then this constraint together with its inverse(s) are used to narrow down the intervals of relavent variables after an integration step. For example, suppose after an integration step, the intervals for variables A, B, and C are A = [5, 6], B = [3, 4], and C = [2.5, 3.5].

Applying interval arithmetic defined in Table 3 to Equation (11), we can determine the range of A by $(B + C)$ as follows: [5.5, 7.5] = [3, 4] + [2.5, 3.5]. Consider the initial interval $A = [5,\ 6]$ from integration, the intersection of these two intervals will be: A= [5, 6] ∩ [5.5, 7.5] = [5.5, 6].

To narrow the rest of the variables the inverse constraints should be used: according to Equation (12), the interval for B should be: [2, 3.5] = [5.5, 6] - [2.5, 3.5]. Again from integration, $B = [3,\ 4]$. Taking the intersection of the intervals gives $B = [3,\ 3.5]$. Similarly, using Equation (13), the interval for C is: [2 , 3] = [5.5, 6] - [3, 3.5], but $C = [2.5,\ 3.5]$ from integration therefore taking the intersection $C = [2.5,\ 3]$.

Finally, intervals for all variables narrowed as much as possible are: [2.5, 3] = [5.5, 6] - [3, 3.5].

After this process the intervals for A, B and C are narrowed as much as possible by reasoning over Constraint (11) (and its two inverses) alone. However, these updated values may result in further narrowing in other constraints; hence the process is repeated until no more changes are made in the whole model or the change is within a given threshold, which is a very small value and determines the simulation precision. Due to the narrowing of intervals using this Inverse Constraint Operations, not much looping of the whole model is required.

The interval narrowing algorithm described above is essentially a Waltz algorithm [Waltz, 1975] applied to interval values. In particular, the soundness and completeness of the Waltz algorithm applied to interval refinement has been extensively studied by Davis [Davis, 1987]. It is noted that in Q3 [Berleant and Kuipers, 1997] the Waltz algorithm was also used to refine interval values. However, Q3 applied the Waltz algorithm to the constraint network composed of constraints instantiated from the model across different time points, one of the reasons for which is to propagate the quantitative information (in the form of intervals) annotated on the given and newly interpolated states throughout the network. While in our interval simulation algorithm the Waltz algorithm is used only at the current time point to ensure the generation of tight interval values for precise estimation of variable values as well as for the integration towards the next time point.

We finally point out that this interval narrowing algorithm is suitable for the situation when there exist time-invariant interval parameters in the model, as described in Equation (3), because the constant intervals for these parameters can be directly processed by the interval arithmetic during interval narrowing. This makes our simulation different from the widely used approach suggested by Lohner [Lohner, 1988] in the NS community, which treats the time-invariant interval parameters as independent state variables and thus increases the complexity of simulation.

5.4 Modes of Simulation

A common problem in interval simulation is that the intervals begin to widen and then eventually become uncontrollable during simulation. This problem is called the "wrapping effect" [Moore, 1966] in NS. In this research we offer different approaches to deal with the interval widening problem, and each approach is termed a *simulation mode* in this paper. For ease of description, the simulation which does not take any additional approach to reduce the spurious behaviours is also considered as a simulation mode: the *Basic Interval Simulation* (BIS) mode.

Basic Interval Simulation

In the *Basic Interval Simulation* (BIS) mode, the previously mentioned three modules are straightforwardly employed to simulate a model. In this sense BIS will demonstrate how the basic non-constructive simulation algorithm is performed.

In the BIS mode, at the beginning of the simulation, users are asked to give the size of the integration step δt and the total simulation time t_{tot}, and therefore the number of simulation steps is given by: $num = t_{tot}/\delta t$.

At the initial time step t_0, given an incomplete initial state, which specifies the initial intervals for some model variables, the interval narrowing algorithm is first used to narrow down as much as possible these initial intervals. Meanwhile, based on the known initial intervals, the interval narrowing algorithm also tries to infer the initial intervals of those variables whose values are not specified. Another function of the interval narrowing algorithm is to check whether the initial state is consistent with the model, and if the initial state is inconsistent, the simulation will not proceed.

During simulation a repository R is maintained to record the history of simulation data, and each element in R is a four-tuple $< Var, Der, [a, b] : t_i >$, where Var is the name of the variable; Der is a non-negative integer which specifies the order of derivative of Var (0 means the magnitude); the third and fourth elements represent the interval value and the time step, respectively.

At each new time step t_i all derivatives of all variables are first integrated, and the relevant intervals used for integration can be retrieved from the repository R. For the Taylor method, only the data at time step t_{i-1} will be retrieved, and the data may include intervals for magnitudes and derivatives of variables. For the Adams-Bashforth method, apart from data at t_{i-1}, data before t_{i-1} will also be retrieved, but only intervals for the first derivatives of relevant variables will be used.

Then in the interval narrowing process, all the intervals for all derivatives are iteratively checked against each constraint of the model. After the interval narrowing process, all the updated intervals are stored in the repository R, which are ready for the simulation at the next time step t_{i+1}. Finally we point out that BIS is complete, as will be discussed in Section 6.1.

Sub-interval Simulation

One way to handle the uncontrollable growth of intervals during simulation is to simulate the model many times but using a smaller sub-interval each time. This approach is called Sub-interval Simulation (SIS) in this paper. The motivation of SIS is that when the initial intervals are very small the simulation will suffer less from the interval divergence over time.

In the SIS mode, the initial interval for each variable/derivative is divided into n equal and non-overlapping sub-intervals. Then for each combination which takes one sub-interval from each initial interval, an initial state is generated and the BIS is used to simulate the model with this initial state. After all possible combinations of sub-intervals have been used for simulation, the final simulation results will be the union of all individual simulations.

As with the Basic Interval Simulation, the Sub-interval Simulation is complete in the sense that it can guaranteed to bound all real solutions, as will be discussed in Section 6.1. Although complete, the Sub-interval Simulation requires a large number of initial states generated for simulation, which becomes a major concern in terms of the computational efficiency. However, it is often the case that not all the generated initial states are consistent with the model, and these states will be discarded during the interval narrowing process, which makes the simulation less expensive.

Monte-Carlo Interval Simulation

Although the sub-interval simulation is complete, its computational cost may increase exponentially with the increase of the number of intervals. Another simulation mode is to use the Monte-Carlo method [Rubinstein, 1981]. That is: instead of exhaustively using all combinations of sub-intervals as in SIS, we only randomly sample a specified number of combinations. This simulation mode is called Monte-Carlo Interval Simulation (MCIS) in this paper.

The sub-intervals used by MCIS can be smaller than those in SIS, which means it can generate a tighter enclosure of the simulation trajectories. Theoretically speaking, MCIS has to sample an infinite number of combinations to bound all real solutions, but as the sample space of MCIS is finite (each interval is divided into a finite number of sub-intervals), compared with the MC method applying to infinite sample space, it is more likely to cover all solutions if the samples are sufficient enough.

Point Simulation

The point simulation samples points (intervals with zero width) rather than sub-intervals from given initial intervals to approximate the solutions. The point simulation is similar to traditional numerical simulation, except that non-constructive simulation is used. The motivation for using points for simulation is to reduce the spurious behaviours, because the trajectories obtained from simulation with point initial conditions are zero width.

Several point simulation methods are developed in this research: (1) the Extreme Point Simulation, in which each initial state is formed by taking the upper or lower bound from each initial interval. Taking the extreme points of each interval gives an approximate range of possible values whilst maintaining an efficient method. This method is sound but incomplete in the sense that the solutions found contains no spurious ones but it may not cover every possible solution.

(2) Regular-spaced Point Simulation: As with the Sub-interval Simulation, the Regular-spaced Point Simulation method takes each interval in the initial state and splits it into a number of states. These states contain a number of regular-spaced points which approximate the interval. This method is theoretically sound and complete when the number of points in each interval tends to infinity. As with the Sub-interval Simulation, this method is exponential in the number of intervals. However, similar to Sub-interval Simulation, it also has the benefit that not every state has to be simulated because some of them will be inconsistent with the model.

(3) Monte-Carlo Point Simulation: For each initial interval, we randomly choose a point within it and thus form an initial state. Then we generate a specified number of such initial states to perform simulation. The advantages of this approach are: using points guarantees that the solution is sound and using Monte-Carlo methods makes the solution tend to be complete and more efficient than Regular-spaced Point Simulation.

A Summary of All Simulation Modes

In this subsection, we proposed several simulation modes to improve the simulation. These simulation modes are classified by two categories: simulation using real intervals and simulation using group of points to approximate the real solutions. We also offer both deterministic and Monte-Carlo approaches to perform the simulation.

In practice, the choices of simulation modes mainly depend on two factors: the requirements of different problems and the computational cost.

6 Theoretical Analysis of the Simulation Algorithm

In this section, we present some theoretical analysis on the completeness, soundness, convergence, and stability of the proposed algorithm. As the algorithm is a collection of integration methods and simulation modes, we have to analyse every combination of simulation modes and integration methods. We first study the completeness and soundness of the algorithm under different simulation modes, then we further analyse the convergence and stability of the algorithm. All proofs of the lemmas and theorems are not presented in this paper due to page limit.

6.1 Completeness and Soundness

The completeness means that the interval simulation algorithm must bound all real solutions. The soundness means that the simulation results should be a subset of the actual solution. As for the completeness, we present the following two theorems:

Theorem 1. *Non-constructive Interval Simulation under the Basic Interval Simulation mode is complete.*

Theorem 2. *Non-constructive Interval Simulation under the Sub-interval Simulation mode is complete.*

BIS is not sound because of the wrapping effect, which is intrinsic to interval arithmetic. Similarly, SIS is also not sound. For the point simulation modes, we give the following lemma and theorem:

Lemma 1. *Non-constructive numerical simulation using the BIS mode is complete and sound.*

Theorem 3. *All Point Simulation modes are sound, but not complete.*

6.2 Convergence and Stability

In this section we discuss the convergence and stability of the algorithm. The convergence and stability analysis presented in this section is influenced by Berleant and Kuipers [Berleant and Kuipers, 1997], who are in turn inspired by Moore [Moore, 1979] and proofs of convergence and stability in numerical simulation [Gear, 1971]. For interval simulation, a simulation algorithm is convergent if at any time point the uncertainty of any variable values is eliminated when the integration step approaches zero and the uncertainty of initial conditions does not exist. Here the uncertainty of a variable value in the context of interval simulation means the width of intervals. For interval simulation, we say a simulation algorithm can achieve $h \to 0$ stability [Berleant and Kuipers, 1997] if at any time point the uncertainty of variable values is bounded by the uncertainty of initial conditions when the integration step h approaches zero.

We first give the following lemma, which shares some similarities with Lemma 1 in [Berleant and Kuipers, 1997].

Lemma 2. *Consider the first order differential equation with a given initial condition Y_0, assume its explicit form is as follows:*

$$Y' = F(Y), \tag{14}$$

where F is an interval valued function of Y, although this explicit form cannot always be obtained as function $\mathcal{F}$ is not always solvable. Suppose $Y(t) \subseteq [lo, hi]$, where $lo, hi \in \mathbb{R}$. We further assume that $F(Y)$ is defined when $Y(t) \subseteq [lo, hi]$, and $F(Y)$ is calculated by using the interval arithmetic defined in Table 3. Suppose the corresponding real rational function of $F(Y)$ is $f(y)$.

Let Y_n be the simulated value for Y at time step n. If the given initial condition $Y_0 \subseteq [lo, hi]$, and the **forward Euler method** *is used, there exists a constant K such that*

$$|Y_n| \leq |Y_0| + Kh \tag{15}$$

In the above operator $|\cdot|$ denotes the width of an interval; h is the integration step size; Y_n is the simulated interval at time step n.

Similarly, if the two-step AB method and Taylor method are used, we have the following two lemmas:

Lemma 3. *If all the assumptions are the same as those made in Lemma 2 except that* **the AB method** *is used, we still have statement (15).*

Lemma 4. *In Lemma 2 if the Taylor method is used, we can still have statement (15).*

Equipped with the above lemmas, we give the theorem of convergence and stability:

Theorem 4. *(Convergence and Stability for a system of ODEs) Consider the following system of first order differential equations:*

$$\vec{\mathcal{F}}(t, \mathbf{Y}, \mathbf{Y}') = 0 \tag{16}$$

where $\mathbf{Y}$ is an interval valued vector, and $\vec{\mathcal{F}}$ is a vector of interval valued functions of $\mathbf{Y}$. Assume for each element of vector $\mathbf{Y}$, $Y_{(i)}(t) \subseteq [lo, hi]$. We further assume that we can solve $\vec{\mathcal{F}}$ in (16) to obtain the explicit form as follows:

$$\mathbf{Y}' = \mathbf{F}(\mathbf{Y}), \tag{17}$$

where $\mathbf{F}$ is a vector of interval valued functions of $\mathbf{Y}$. Suppose for each element of $\mathbf{Y}$, $Y_{(j)}(t) \subseteq [lo, hi]$, where $lo, hi \in \mathbb{R}$. We further assume that $\mathbf{F}(\mathbf{Y})$ is defined when each $Y_j(t) \subseteq [lo, hi]$, and for each element of $\mathbf{F}(\mathbf{Y})$, $F_{(j)}$ is calculated by using the interval arithmetic defined in Table 3. Suppose the corresponding real rational function of $F(Y)$ is $f(y)$.

Let $\mathbf{Y_n}$ be the simulated value for $\mathbf{Y}$ at time step n. If for each element of the given initial condition $\mathbf{Y_0}$, $Y_{0j} \subseteq [lo, hi]$, and the **Tayor or AB integration methods** *are used, there exists a constant K such that*

$$||\mathbf{Y_n}|| \leq ||\mathbf{Y_0}|| + Kh \tag{18}$$

In the above, operator $||\cdot||$ denotes the norm of an interval valued vector such that for an interval valued vector $\mathbf{A} = \{A_1, A_2, \cdot\cdot, A_n\}$, $||\mathbf{A}|| = max(|A_1|, |A_2|, \cdot\cdot, |A_n|)$, where $|\cdot|$ denotes the width of an interval; h is the integration step size; $||\mathbf{Y_n}||$ is the simulated interval vector at time step n.

It is noted that Theorem 4 can be extended to higher order systems as any such systems can be reduced to first order systems. From statement (18), we see that given a precise initial condition $||\mathbf{Y_0}|| = 0$, for any fixed simulation time $t = nh$, when $h \to 0$, we will have $||\mathbf{Y_n}|| \to 0$. This means that the algorithm is *convergent* when the simulation step approaches zero. According to the $h \to 0$ stability defined in [Berleant and Kuipers, 1997] and shown below:

$$||\mathbf{Y_n}|| \leq K||\mathbf{Y_0}||, \tag{19}$$

where K is a constant, we can see that our proposed simulation algorithm possesses the $h \to 0$ stability from Theorem 4.

In Theorem 4 we assume that (16) can be solved to obtain its explicit form (17). This actually indicates that the underlying model is a system of ODEs. As we know that (16) could also be a system of DAEs, we give the following theorem:

Theorem 5. *(Convergence and Stability for a system of DAEs) Suppose by solving (16), we can only obtain the following form:*

$$\mathbf{X}' = \mathbf{F}(\mathbf{X}, \mathbf{Z}), \tag{20}$$

$$\mathbf{0} = \mathbf{G}(\mathbf{X}, \mathbf{Z}). \tag{21}$$

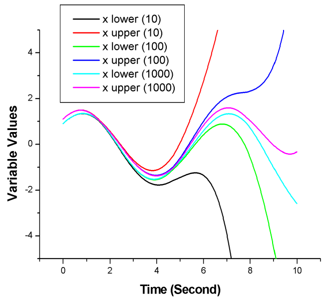
In the above $\mathbf{F}$ and $\mathbf{G}$ are two vectors of interval valued functions. Vector $\mathbf{Y}$ in (16) can be formed by combining vectors $\mathbf{X}$ and $\mathbf{Z}$ together. (This means that the underlying model is a system of DAEs.)

If we assume that $\mathbf{F}$ and $\mathbf{G}$ are defined when each $Y_j(t) \subseteq [lo, hi]$, and the other assumptions are the same as those in Theorem 4, there exist three constants K_1, K_2, and K_2 such that

$$||\mathbf{X_n}|| \leq ||\mathbf{X_0}|| + K_1 h \tag{22}$$

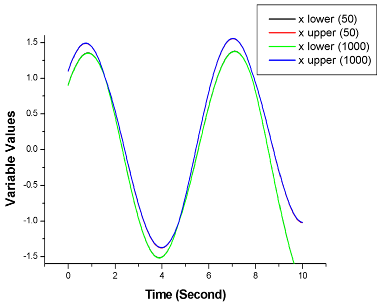
$$||\mathbf{Z_n}|| \leq K_2||\mathbf{Z_0}|| + K_3 h \tag{23}$$

From this theorem we can conclude that the convergence and stability can still be achieved when dealing with DAEs. Finally, from Theorems 4 and 5 we see that in nonconstructive interval simulation the uncertainty of simulation results measured by the norm of vectors at each simulation step is determined by two factors: the uncertainty of initial



"x lower" and "x upper" mean the lower and upper bounds of x, respectively. Numbers in the brackets indicate the numbers of sub-intervals being used.

Figure 1: Sub-interval Simulation



"x lower" and "x upper" mean the lower and upper bounds of x, respectively. Numbers in the bracket indicate the numbers of Monte-Carlo Sample Intervals being used. In this figure as the results of simulation using 50 samples and those using 1000 samples largely overlapped, only two curves can be clearly seen.

Figure 2: Monte-Carlo Interval Simulation

conditions and the size of the simulation step. To reduce the uncertainty, we can either use a smaller simulation step or split initial intervals into several subintervals. This justifies the use of all the simulation modes in addition to BIS.

7 Experiments

In this section we only report part of the experimental results due to page limit, and readers are referred to [Pang *et al.*, 2012] or the supplementary material for a full experiment report.

The model for the spring-mass system is given as: $x'' = F - kx$, where F is the constant external force, k is a constant parameter, and x is the displacement of the mass with respect to the equilibrium position. The initial condition is set as follows: $x = [0.9, 1.1]$, $x' = [1, 1]$, $F = [0, 0]$. The sub-interval simulation and Monte-Carlo Interval Simulation results are shown in Figure 1 and Figure 2, respectively.

The Van der Pol oscillator model is given as $x'' = -P(x^2 - 1)x' - Qx$, where P and Q are two parameters. The Regular-spaced Point simulation for this model is given in Figure 3, where the initial condition is x, $x' = [0.5, 1.5]$ and $P, Q = [1, 1]$. The Monte-Carlo Points simulation on this model is given in Figure 4, where the initial condition is

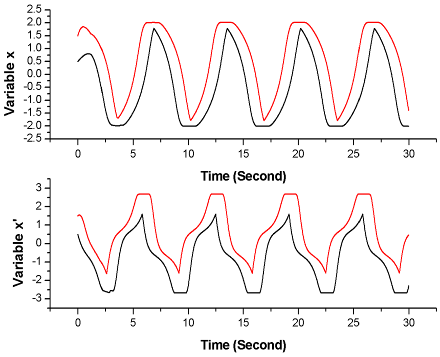


Figure 3: Regular-spaced Point Simulation of the Van der Pol Oscillator (20 Points Per Interval)

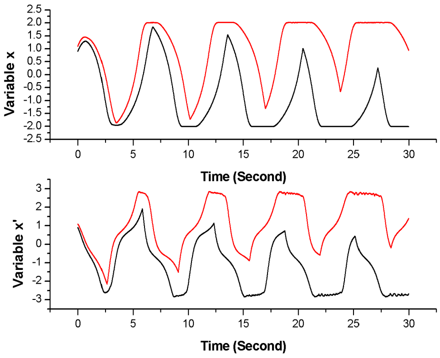


Figure 4: Monte-Carlo Points Simulation of the Van der Pol Oscillator with All Variables taking Interval Values

$x, x', P, Q = [0.9, 1.1]$ (all initial values and parameter values are intervals).

8 Conclusions and Future Work

In this paper we have presented a novel non-constructive interval approach for the simulation of dynamic systems. We established our novel non-constructive interval simulation approach by (1) recasting existing NS integration methods which are feasible for non-constructive simulation, (2) providing an iterative interval narrowing algorithm to deal with the interval widening effect, and (3) offering several simulation modes to meet different requirements.

In the future more simulation modes will be investigated and implemented so that we can better sample the sub-intervals or points from initial intervals of variables and parameters. Finally, we expect that the research results presented in this paper will contribute to both the NS and QR communities, and we foresee the non-constructive approach as a fruitful research direction for simulation at both quantitative and semi-quantitative levels.

References

[Angermann, 2011] Lutz Angermann, editor. *Numerical Simulations - Applications, Examples and Theory*. InTech, 2011.

[Arfken *et al.*, 2005] George B. Arfken, Hans J. Weber, and Frank E. Harris. *Mathematical Methods for Physicists, Sixth Edition: A Comprehensive Guide*, chapter 5.6, pages 352–362. Academic Press, 2005.

[Ascher and Petzold, 1998] Uri M. Ascher and Linda Ruth Petzold. *Computer methods for ordinary differential equations and differential-algebraic equations*, chapter 3, pages 37–67. SIAM, 1998.

[Berleant and Kuipers, 1997] Daniel Berleant and Benjamin Kuipers. Qualitative and quantitative simulation: bridging the gap. *Artificial Intelligence*, 95(2):215–255, 1997.

[Bruce and Coghill, 2005] A. M. Bruce and G. M. Coghill. Parallel fuzzy qualitative reasoning. In *Proceedings of the 19th International Workshop on Qualitative Reasoning*, pages 110–116, Graz, Austria, 2005.

[Burden and Faires, 2000] Richard L. Burden and J. Douglas Faires. *Numerical Analysis (7th Edition)*, chapter 2, pages 56–66. Brooks Cole, 2000, December 2000.

[Butcher, 2003a] John C. Butcher. *Numerical Methods for Ordinary Differential Equations*, pages 97–106. John Wiley, 2003.

[Butcher, 2003b] John C. Butcher. *Numerical Methods for Ordinary Differential Equations*, page 103. John Wiley, 2003.

[Butcher, 2008] John Charles Butcher. *Numerical methods for ordinary differential equations (2nd Edition)*, chapter 2, pages 93–103. John Wiley and Sons, 2008.

[Cellier, 1991a] Francois E. Cellier. *Continuous System Modeling*, chapter 5, pages 167–169. Springer-Verlag, 1991.

[Cellier, 1991b] Francois E. Cellier. *Continuous System Modeling*, chapter 3, pages 66–68. Springer-Verlag, 1991.

[Coghill and Chantler, 1999] G. M. Coghill and M. J. Chantler. Constructive and non-constructive asynchronous qualitative simulation. In *Proceedings of the 13th Inernational Workshop on Qualitative Reasoning*, pages 51–61, Loch Awe, Scotland, June 1999.

[Coghill, 1996] George M. Coghill. *Mycroft: A Framework for Constraint based Fuzzy Qualitative Reasoning*. PhD thesis, Heriot-Watt University, September 1996.

[Davis, 1987] Ernest Davis. Constraint propagation with interval labels. *Artificial Intelligence*, 32(3):281–331, July 1987.

[Gear, 1971] C. William Gear. *Numerical Initial Value Problems in Ordinary Differential Equations*. Prentice Hall, 1971.

[Hairer *et al.*, 1993] Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*, volume 8 of *Springer Series in Computational Mathematics*, chapter III.1, pages 356–366. Springer, 1993.

[Kuipers and Berleant, 1988] B. J. Kuipers and D. Berleant. Using incomplete quantitative information in qualitative reasoning. In *Proceedings of the National Conference on Artificial Intelligence (AAAI-88)*, pages 324–329, Los Altos, CA, 1988. Morgan Kaufman.

[Kuipers, 1994] Benjamin Kuipers. *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge*. MIT Press, Cambridge, MA, 1994.

[Lohner, 1988] Rudolf Lohner. *Enclosing the solutions of ordinary initial and boundary value problems (in German)*. Ph.D. thesis, Universität Fridericiana Karlsruhe, Karlsruhe, Germany, 1988.

[Mathews and Fink, 2004] John H. Mathews and Kurtis K. Fink. *Numerical Methods Using MATLAB (4th Edition)*, chapter 9.6, pages 505–508. Prentice-Hall Inc., 2004.

[Moore, 1966] Ramon E. Moore. *Interval Analysis*. Prentice-Hall, Englewood Cliff, New Jersey, 1966.

[Moore, 1979] Ramn E. Moore. *Methods and applications of interval analysis*, chapter 8, pages 93–99. SIAM, 1979.

[Nedialkov and Pryce, 2005] Nedialko S. Nedialkov and John D. Pryce. Solving differential algebraic equations by Taylor series (I): Computing Taylor coefficients. *BIT Numerical Mathematics*, 45(3):561–591, 2005.

[Nedialkov *et al.*, 1999] N. S. Nedialkov, K. R. Jackson, and G. F. Corliss. Validated solutions of initial value problems for ordinary differential equations. *Applied Mathematics and Computation*, 105(1):21–68, October 1999.

[Pang *et al.*, 2012] Wei Pang, G M. Coghill, and Alan M. Bruce. Non-constructive interval simulation of dynamic systems. Technique Report ABDNCS1202, Department of Computing Science, University of Aberdeen, 2012.

[Press *et al.*, 1992] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in FORTRAN: The Art of Scientific Computing (2nd Edition)*, chapter 16, pages 740–744. Cambridge University Press, 1992.

[Pryce, 2001] J. D. Pryce. A simple structural analysis method for DAEs. *BIT Numerical Mathematics*, 41(2):364–394, 2001.

[Rubinstein, 1981] Reuven Rubinstein. *Simulation and the Monte Carlo Method*. Wiley, 1981.

[Shen and Leitch, 1993] Qiang Shen and Roy Leitch. Fuzzy qualitative simulation. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(4):1038–1061, 1993.

[Waltz, 1975] David Waltz. Understanding line drawings of scenes with shadows. In Patrick Winston, editor, *The Psychology of Computer Vision*, pages 19–91. McGraw-Hill, 1975.

[Wiegand, 1991] M. Wiegand. *Constructive Qualitative Simulation of Continuous Dynamic Systems*. PhD thesis, Heriot-Watt university, May 1991.

Helping Investors to assess new ventures proposals by a linguistic Multi-Criteria Decision tool

Xari Rovira, David Riu, Ferran Gionés, Núria Agell, Mònica Casabayó

Abstract

Capturing experts knowledge when evaluating new venture proposals in investment decisions is an active and ongoing area of research. Determining the agreement among experts on the importance of the variables involved in these decisions is crucial. This study presents a new approach to better understanding the multi-criteria decision making process of expert investors that is based on a combination of the Order Preference by Similarity to an Ideal Solution method (TOPSIS) and a measure of consensus. A real case has been conducted among 28 experts to understand the relevant variables by obtaining their opinions by means of hesitant linguistic terms. Results provide not only a ranked list of variables obtained from the opinions of the group of experts but also a measure of the consensus reached on their importance.

1 Introduction

Much has been written since the early 1970s on the factors influencing investors when making investment decisions about entrepreneurial ventures – and the criteria investors use to assess proposals continue to be of interest to scholars and practitioners [Franke *et al.*, 2008]. Specific areas of interest include: examining the complexity of the investment decision making process and the variables that define the criteria used [Carlos Nunes *et al.*, 2014]; how the information is gathered and evaluated [Zacharakis and Shepherd, 2001]; defining highly uncertain environments [Huang and Pearce, 2015]; and how investors deal with uncertainty [Dimov and De Clercq, 2006; Bertoni *et al.*, 2013].

Given the complexity of this decision-making process, several authors have focused on listing the numerous factors that experts consider. For example, Van Osnabrugge [Van Osnabrugge, 2000] outlines more than 20 traits that possible investors examine before making a decision. Moreover, Streletzki and Schulte [Streletzki and Schulte, 2013] provide a distinction between both types of criteria and their degrees of importance for better evaluating venture proposals. For a comprehensive review of variables, see the ones identified by Nunes et al. [Carlos Nunes *et al.*, 2014]. as well as the review by Maxwell et al. And, see [Maxwell *et al.*, 2011] for building the reference background to explore how investors relate to these criteria. These are the theoretical bases of our real case experiment.

When capturing the multiple factors that influence decisions, existing approaches fail to reduce the mentioned degree of uncertainty. To overcome this situation, firstly, behavioural decision theory has shown that entrepreneurs [Koellinger *et al.*, 2007] and independent inventors [Åstebro *et al.*, 2007] fall to possible biases and heuristics that affect individual decision-making. Secondly, although there is not a big gap between the relevant factors to consider, there is not a consensus when ranking them as different experts may prioritise the factors within a different way. Maxwell et al [Maxwell *et al.*, 2011] emphasises the importance of group consensus among investors, despite being an individual decision, when agreeing on the potential of an opportunity. Thus, collecting and interpreting the emerging knowledge on how experts make their investment decision and providing the investors community with an aggregate consensus view is also a relevant challenge for academics and practitioners in the field.

Furthermore, there is ongoing research about the most appropriate method to solve the investors decision making problem. A close examination of the literature by Petty and Gruber [Petty and Gruber, 2011] reveals that the majority of studies rely exclusively on post-hoc research methods that typically depend on cross-sectional survey or interview data where investors are asked to list and rank their own evaluation criteria. Studies using experimental approaches are also found in the literature. These approaches are based on data gathered in real.time from a group of investors using a conjoint analysis. However, despite the ability of experimental studies to improve the validity of research and to offer new insights on investors decision making, these approaches also suffer from limitations due to the lack of flexibility in the variables expression, as explained in [Gustafsson *et al.*, 2007].

We propose a method based on hesitant fuzzy linguistic terms sets [Rodríguez *et al.*, 2012; Montserrat-Adell *et al.*, 2017a] and Fuzzy TOPSIS (Preference by Similarity to an Ideal Solution)[Chen, 2000; Kahraman, 2008] to help understand how investors complete the evaluation of fuzzy opinions for new ventures proposals. In doing so, a consensus group measure based on hesitant fuzzy linguistic terms [Montserrat-Adell *et al.*, 2017b] and an extension of the TOPSIS method

[Afsordegan *et al.*, 2016] are considered.

The real case included in this paper takes into account the opinion of a group of investment experts to assess different variables by means of linguistic terms with different levels of precision. Results offer a ranked list of variables obtained from a group of experts opinion and a measure of the consensus reached.

The paper is organized as follows. Section 2 establishes the theoretical framework regarding the consensus measure and the ranking method that is applied in the paper. In Section 3 a real-case application analyzing how investors asses entrepreneurship projects is detailed along with research findings and results. Finally, the conclusions and directions for further research are explained in Section 4.

2 Methodology

This section describes the mathematical structures and tools needed for the development. Firstly the preliminary concepts related to linguistic descriptions, including a distance measure, are described [Agell *et al.*, 2015; Montserrat-Adell *et al.*, 2017a]; secondly, the measure of consensus [Montserrat-Adell *et al.*, 2017b]; and finally a methodology for multi-criteria decision making based on TOPSIS [Afsordegan *et al.*, 2016; Hwang and Yoon, 1981]. These concepts will enable us to define a ranking method that gives an ordered set of alternatives, and a consensus value to measure agreement among decision-makers.

2.1 Preliminaries

Let $\mathcal{S}$ be a finite total ordered set of linguistic terms $\mathcal{S} = \{a_1, \ldots, a_n\}$ with $a_1 < \ldots < a_n$, we consider a hesitant fuzzy linguistic term set (HFLTS) over $\mathcal{S}$ as a subset of consecutive linguistic terms of $\mathcal{S}$, i.e. $\{x \in \mathcal{S} \,|\, a_i \leq x \leq a_j\}$, for some $i, j \in \{1, \ldots, n\}$ with $i \leq j$ [Agell *et al.*, 2015]. We note the set $\{x \in \mathcal{S} \,|\, a_i \leq x \leq a_j\}$ as $[a_i, a_j]$ if $i \leq j$ or $\{a_i\}$ if $i = j$. In addition, $\mathcal{H}_\mathcal{S}$ is the set of all the possible HFLTSs over $\mathcal{S}$, being $\mathcal{H}^*_\mathcal{S} = \mathcal{H}_\mathcal{S} - \{\emptyset\}$. This set is extended in [Montserrat-Adell *et al.*, 2017a] to $\overline{\mathcal{H}_\mathcal{S}}$, including the concept of *negative* HFLTSs as the set of gaps between pairs of HFLTSs. In this framework, we consider the operators *extended connected union* and *extended intersection* in $\overline{\mathcal{H}_\mathcal{S}}$. Given $H_1, H_2 \in \overline{\mathcal{H}_\mathcal{S}}$, the *extended intersection* of H_1 and H_2, $H_1 \sqcap H_2$, is the largest element in $\overline{\mathcal{H}_\mathcal{S}}$ that is contained in H_1 and H_2 and the *extended connected union* of H_1 and H_2, $H_1 \sqcup H_2$, is the least element in $\overline{\mathcal{H}_\mathcal{S}}$ that contains H_1 and H_2. A detailed explanation of this concepts can be found in [Montserrat-Adell *et al.*, 2017a].

Finally, we consider in $\overline{\mathcal{H}_\mathcal{S}}$ the distance between HFLTSs defined in [Montserrat-Adell *et al.*, 2018] based on the *width* of H, $\mathcal{W}(H)$, defined as the cardinal of H if $H \in \overline{\mathcal{H}_\mathcal{S}}$ or $-card(-H)$ if H is a negative HFLTS. Given $H_1, H_2 \in \overline{\mathcal{H}_\mathcal{S}}$, we consider:

$$D(H_1, H_2) = \mathcal{W}(H_1 \sqcup H_2) - \mathcal{W}(H_1 \sqcap H_2). \qquad (1)$$

The following example illustrates the previous concepts.

Example 2.1 ([Montserrat-Adell *et al.*, 2018]) *Let a_1 = not important, a_2 = slightly important, a_3 = moderately important, a_4 = important and a_5 = very important, be five linguistic labels defining the set $\mathcal{S} = \{a_1, a_2, a_3, a_4, a_5\}$. Note that as can be seen in Figure 1 the values $\mathcal{S} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$ are the landmarks defining the initial labels.*

Three possible assessments in $\mathcal{H}_\mathcal{S}$ are A = "below moderately important", B = "very important" and C = "neither very important nor not important" and their corresponding HFLTS by means of $\mathcal{S}$ are $H_A = [a_1, a_2]$, $H_B = \{a_5\}$ and $H_C = [a_2, a_4]$ respectively.

In this case, if we consider, for instance, the extended connected union and extended intersection of H_A and H_B, we obtain, as can be seen in Figure 1, $H_A \sqcup H_B = [a_1, a_5]$ and, $H_A \sqcap H_B = -[a_3, a_4]$, because they are respectively, the largest element in the $\overline{\mathcal{H}_\mathcal{S}}$ lattice that is contained in H_A and H_B, and the least element in $\overline{\mathcal{H}_\mathcal{S}}$ that contains H_A and H_B.

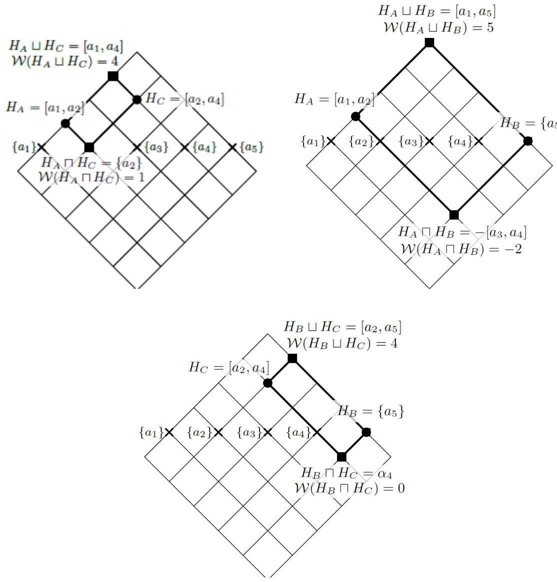


Figure 1: Extended connected union and extended intersection of two HFLTSs.

According to these results, and considering the given definition of distance, $D(H_A, H_B) = 5 - (-2) = 7$, $D(H_A, H_C) = 4 - 1 = 3$ and $D(H_B, H_C) = 4 - 0 = 4$.

2.2 Consensus process

Consensus measures are considered to capture the level of agreement between a set of decision-makers regarding a given set of alternatives [Herrera-Viedma *et al.*, 2002; Rodríguez and Martínez, 2015]. The use of hesitant fuzzy linguistic terms sets in consensus measurement enables us to capture the imprecision included in the opinion of decision-makers. In this paper, we consider the consensus approach defined in Montserrat-Adell et al. [Montserrat-Adell *et al.*, 2018] in which overlapping opinions among decision-makers are considered, as well as the order of magnitude of the existing discrepancies among them. This measure is based on the distance defined in 1. In this section, we introduce the concept of *central opinion* (or centroid) to define the consensus degree of a group G of k decision-makers $G = \{g_1, \ldots, g_k\}$ on a specific alternative. To this end, let $H_1, \ldots, H_k$ be the opinion of the decision-makers when assessing this alternative. The *central opinion* of G on this alternative is:

$$C = arg \min_{H \in \mathcal{H}_{\mathcal{S}}^*} \sum_{i=1}^{k} D(H, H_i). \quad (2)$$

The consensus degree of G is then defined as a normalization of the addition of distances between the centoid of G and the opinion of each decision-maker:

$$\delta(G) = 1 - \frac{\sum_{i=1}^{k} D(C, H_i)}{k \cdot (n-1)}. \quad (3)$$

The objective of this consensus degree is to measure the level of agreement within the group of decision-makers. This value will be considered as a validity measure for the global opinion of the group.

2.3 Ranking process

Several decision making approaches and applications have been extended by using HFLTSs [Fahmi *et al.*, 2016; Rodríguez *et al.*, 2013; Rodríguez *et al.*, 2014; Rodríguez *et al.*, 2016]. In this section, we present a modified version of the TOPSIS method based on the concepts of HFLTS introduced in Subsection 2.1. The Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS), developed by Hwang and Yoon [Hwang and Yoon, 1981], is one of the most well-known distance-based approaches for such decision making. TOPSIS method ranks the alternatives with respect to their distance from the positive and negative ideal solutions. This approach uses value assessments expressed through crisp values. However, in real situations when analyzing and assessing different alternatives, the opinions given by decision-makers are imprecise and better captured using HFLTS.

TOPSIS method considers the distance between alternatives and certain target points, named ideal positive solution and ideal negative solution which model the best and the worst alternatives, respectively. The values of the distances to this target points are used to rank the given alternatives. To consider TOPSIS method in a HFLTS framework we propose a new approach, Hesitant Fuzzy TOPSIS (HFTOPSIS), where decision-makers can use HFLTS to asses alternatives. Given the group of decision-makers G and a set of r alternatives, we define the *Ideal Negative Solution* $I^- = (b_1^-, \ldots, b_k^-) \in (\mathcal{H}_{\mathcal{S}}^*)^k$ and the *Ideal Positive Solution* $I^+ = (b_1^+, \ldots, b_k^+) \in (\mathcal{H}_{\mathcal{S}}^*)^k$ being $b_i^- = inf\{H_i^1, \ldots, H_i^r\}$ and $b_i^+ = sup\{H_i^1, \ldots, H_i^r\}$, where $H_i^j \in \mathcal{H}_{\mathcal{S}}^*$ is the opinion of decision-maker g_i on the alternative j, $1 \leq j \leq r$.

Distances between the vector of all the assessments on a specific alternative j with respect to both the *Ideal Negative Solution* and the *Ideal Positive Solution* are defined as:

$$d_-^j(I^-, (H_1^j, \ldots, H_k^j) = \sum_{i=1}^{k} D(I_i^-, H_i^j)). \quad (4)$$

$$d_+^j(I^+, (H_1^j, \ldots, H_k^j) = \sum_{i=1}^{k} D(I_i^+, H_i^j)). \quad (5)$$

where D is the distance defined in 1.

Finally, the *Closeness Coefficient* of each alternative j is obtained as:

$$CC_j = \frac{d_+^j}{d_-^j + d_+^j}, \quad (6)$$

and the alternatives are ranked according to the increasing order of CC_j values.

3 Real case application

The EIX Investors Network within La Salle Technova is a community of private investors with considerable experience and track record in financing venture proposals. Highly motivated and active in the evaluation of proposals, this investor community connects entrepreneurs in the field of innovation and technology with investors to facilitate the access of high-tech companies to capital (mainly in the initial phases). Technova Network has been active for over ten years, and was awarded the Best Business Angels Network Prize by the Spanish Association of Business Angels (AEBAN) in 2017.

The main goal of this real case study is to capture the most important factors (or criteria) when assessing an investment and to measure the level of consensus. Not only this information will provide a better understanding of the complex decision process, but also will be very valuable for entrepreneurs to improve on what is really important when looking for financing their venture through investors. To this end, our data comes from the opinions of 35 expert investors from La Salle Technova Network.

3.1 Data Set

For collecting the data, 35 expert investors were given a list of 21 factors (alternatives) grouped by category (see Figure 2) and asked to give their opinion about the importance of each factor when assessing a venture proposal. The importance of each factor was measured on a five-point Likert scale. Respondents could answer with hesitancy, i.e., multiple connected responses, from *not important* to *very important*.

Finally, 28 complete responses (80% response rate) were gathered. The responses were received between March and May 2016. Regarding the investors profiles: 9 of 28 (32%) invest on average less than 50K EUR per funding round; 13 of them (47%) invests between 50K and 150K, and 6 of them (21%) over 150K per round. Most of the investors in the final sample (71%) make fewer than four investments rounds per year, eight of them (29%) make more than four investment rounds per year.

3.2 Results

Following the consensus degree and the FHTOPSIS method presented in 2, both the index of agreement and the ranking of the importance of the factors were obtained (see Figure 3).

Figure 2: Factors grouped by category

Product or service	V1 Interesting product/service (adoption)
	V2 Innovative (high quality) product/service (adoption)
	V3 Technology development risk (product status)
	V4 Protectable IP (Protectable)
Target market	V5 Market validation (Customer engagement).
	V6 Market entry (route to market).
	V7 Supply and distribution partners (route to market).
	V8 Market size (Market potential)
	V9 Market growth / Growth Potential (Market potential).
	V10 Market competitiveness (Market potential).
Entrepreneur and team	V11 Entrepreneur's Industry experience (relevant experience).
	V12 Entrepreneur's Management ability (relevant experience).
	V13 Entrepreneur's Passion
	V14 Entrepreneur's Integrity
	V15 Team experience (relevant experience).
	V16 Team record (relevant experience).
Financial	V17 Cash Flow (Financial Model)
	V18 Expectations / Profitability (Financial Model)
	V19 Realistic Forecast (Financial Model)
	V20 Investment plan
	V21 Co-investments / investors

This figure exhibits then not only the position of each variable in the ranking but also its validation through the consensus degree. The category of each factor has been added for further discussion. Results show primarily, the two factors V14: Entrepreneurs Integrity and V12: Entrepreneur Management ability are demonstrated to be the most important variables for the respondents. Moreover, these two most important factors are also validated as show to get the biggest consensus. Contrary, V4: Protectable IP, shows to be considered the less important factor between the respondents when assessing a venture proposal. However, it is also important to note that V4 also displays the less consensus index between the investors. Despite the 2 most important factors and the last important factor are clearly identified, results ratify the complexity of the decision found in the extended literature, as there are 7 variables that shows a close importance.

3.3 Research findings and managerial implications

The HFTOPSIS gives the literature a new approach to help investors to assess venture proposals. Not only because it provides a ranking of the most important factors to be considered when assessing a new venture but also a consensus index to help validating the consensus between the experts. Results ratify the complexity of the decision, as several factors are considered. Furthermore, In addition, most factors show a similar importance rate which makes the decision of the manager still difficult. Figure 4 and Figure 5 illustrated the importance of the factors and the consensus of the investors according the belonging category. As Figure 4 shows, the experts consider the Entrepreneur category as the most important category when assessing the venture proposal. Financial information seems to be the last important one. However, Product and Service category shows the wider dispersion.

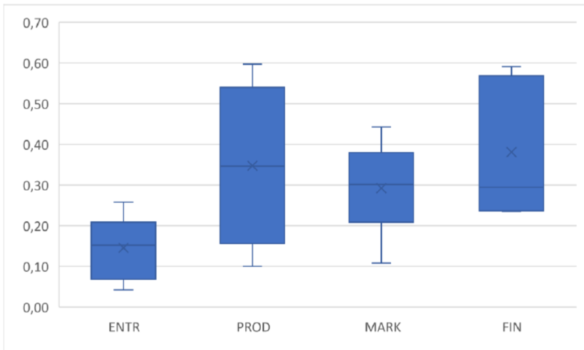
Figure 5 illustrates the consensus between the respondents by category. It is not surprisingly that Entrepreneur category shows the highest consensus and Product category the lowest and with the highest dispersion.

Therefore, entrepreneurs need to know that they are the most important factors when being considered by the investors.

Figure 3: Factors ordered by their relative proximity to the ideal maximum

	VARIABLE	RELATIVE PROXIMITY TO IDEAL MAX (^2)	CONSENSUS
ENTR	V14 Entrepreneur's Integrity	0,04	0,86
ENTR	V12 Entrepreneur's Management ability (relevant experience).	0,08	0,77
PROD	V1 Interesting product/service (adoption)	0,10	0,73
MARK	V5 Market validation (Customer engagement).	0,11	0,71
ENTR	V13 Entrepreneur's Passion	0,14	0,70
ENTR	V11 Entrepreneur's Industry experience (relevant experience).	0,16	0,76
ENTR	V15 Team experience (relevant experience).	0,19	0,67
FIN	V21 Co-investments / investors	0,236	0,66
FIN	V18 Expectations / Profitability (Financial Model)	0,237	0,65
MARK	V8 Market size (Market potential)	0,242	0,66
ENTR	V16 Team record (relevant experience).	0,258	0,71
MARK	V6 Market entry (route to market).	0,291	0,63
FIN	V17 Cash Flow (Financial Model)	0,294	0,61
MARK	V10 Market competitiveness (Market potential).	0,312	0,64
PROD	V2 Innovative (high quality) product/service (adoption)	0,324	0,70
MARK	V9 Market growth / Growth Potential (Market potential).	0,359	0,63
PROD	V3 Technology development risk (product status)	0,368	0,66
MARK	V7 Supply and distribution partners (route to market).	0,443	0,55
FIN	V20 Investment plan	0,545	0,61
FIN	V19 Realistic Forecast (Financial Model)	0,591	0,68
PROD	V4 Protectable IP (Protectable)	0,597	0,55

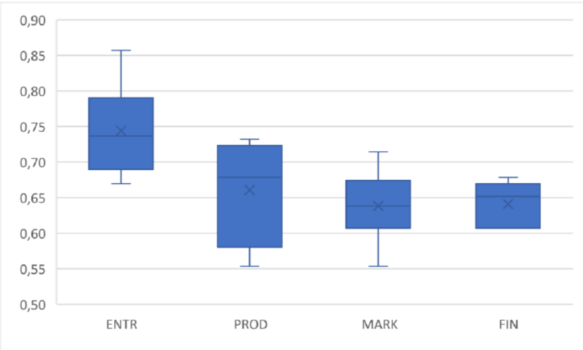
Figure 4: Importance by Category



4 Conclusions and further study

This paper presents a methodology based on a combination of two techniques. On the one hand, a new approach of the well-known multi-criteria decision analysis TOPSIS method is developed when dealing with hesitant assessments of decision-makers to rank a set of alternatives. On the other hand, we introduce a measure of the consensus degree among the group of decision-makers, that allows us to validate the agreement on the importance of the different alternatives. The main difference between the proposed methodology and existing ones

Figure 5: Consensus by Category



is its capability to aggregate opinions on different levels of precision and simultaneously asses its validity by means of a consensus measure.

The proposed method is applied to a real case on evaluation of investment projects. Applying the methodology to data coming from the opinion of a group of experts on the importance of different factors when making the investment decision, will help investors to assess venture proposals. The results support significant contribution to the area of venture capital.

As future work, we are interested in developing a software tool able to automatically evaluate an investment project, combining the importance of the factors, to determine its interest for a venture investor. Further experimentation is planned in order to assess the weights of the different factors, according to the importance of each criteria from the investor point of view. A learning system will be considered to define an aggregation function to capture experts opinions regarding different types of projects. To conclude, let us remark that the methodology introduced in this paper has a wide range of potential applications in management, including evaluation or accreditation processes in different domains.

Acknowledgments

This work was partially supported by the INVITE Research Project (TIN2010-20966-C02-01), funded by the Spanish Ministry of Economy and Competitiveness.

References

[Afsordegan *et al.*, 2016] Arayeh Afsordegan, M Sánchez, N Agell, Siamak Zahedi, and LV Cremades. Decision making under uncertainty using a qualitative topsis method for selecting sustainable energy alternatives. *International journal of environmental science and technology*, 13(6):1419–1432, 2016.

[Agell *et al.*, 2015] N Agell, M Sánchez, F. Prats, and F. J. Ruiz. Computing distances between decision makers using hesitant fuzzy linguistic term sets. *Frontiers in Artificial Intelligence and Applications*, 277:71–79, 2015.

[Åstebro *et al.*, 2007] Thomas Åstebro, Scott A Jeffrey, and Gordon K Adomdza. Inventor perseverance after being told to quit: The role of cognitive biases. *Journal of behavioral decision making*, 20(3):253–272, 2007.

[Bertoni *et al.*, 2013] Fabio Bertoni, Massimo G Colombo, and Luca Grilli. Venture capital investor type and the growth mode of new technology-based firms. *Small Business Economics*, 40(3):527–552, 2013.

[Carlos Nunes *et al.*, 2014] José Carlos Nunes, Elisabete Gomes Santana Félix, and Cesaltina Pacheco Pires. Which criteria matter most in the evaluation of venture capital investments? *Journal of Small Business and Enterprise Development*, 21(3):505–527, 2014.

[Chen, 2000] Chen-Tung Chen. Extensions of the topsis for group decision-making under fuzzy environment. *Fuzzy sets and systems*, 114(1):1–9, 2000.

[Dimov and De Clercq, 2006] Dimo Dimov and Dirk De Clercq. Venture capital investment strategy and portfolio failure rate: A longitudinal study. *Entrepreneurship Theory and Practice*, 30(2):207–223, 2006.

[Fahmi *et al.*, 2016] A. Fahmi, C. Kahraman, and Ü. Bilen. Electre i method using hesitant linguistic term sets: An application to supplier selection. *International Journal of Computational Intelligence Systems*, 9(1):153–167, 2016.

[Franke *et al.*, 2008] Nikolaus Franke, Marc Gruber, Dietmar Harhoff, and Joachim Henkel. Venture capitalists' evaluations of start-up teams: Trade-offs, knock-out criteria, and the impact of vc experience. *Entrepreneurship Theory and Practice*, 32(3):459–483, 2008.

[Gustafsson *et al.*, 2007] Anders Gustafsson, Andreas Herrmann, and Frank Huber. *Conjoint measurement: methods and applications*. Springer Science & Business Media, 2007.

[Herrera-Viedma *et al.*, 2002] Enrique Herrera-Viedma, Francisco Herrera, and Francisco Chiclana. A consensus model for multiperson decision making with different preference structures. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 32(3):394–402, 2002.

[Huang and Pearce, 2015] Laura Huang and Jone L Pearce. Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Science Quarterly*, 60(4):634–670, 2015.

[Hwang and Yoon, 1981] Ching-Lai Hwang and Kwangsun Yoon. Multiple criteria decision making. *Lecture Notes in Economics and Mathematical Systems*, 186:58–191, 1981.

[Kahraman, 2008] Cengiz Kahraman. *Fuzzy multi-criteria decision making: theory and applications with recent developments*, volume 16. Springer Science & Business Media, 2008.

[Koellinger *et al.*, 2007] Philipp Koellinger, Maria Minniti, and Christian Schade. i think i can, i think i can: Overconfidence and entrepreneurial behavior. *Journal of economic psychology*, 28(4):502–527, 2007.

[Maxwell *et al.*, 2011] Andrew L Maxwell, Scott A Jeffrey, and Moren Lévesque. Business angel early stage decision making. *Journal of Business Venturing*, 26(2):212–225, 2011.

[Montserrat-Adell *et al.*, 2017a] Jordi Montserrat-Adell, Núria Agell, Mónica Sánchez, Francesc Prats, and Francisco Javier Ruiz. Modeling group assessments by means of hesitant fuzzy linguistic term sets. *Journal of Applied Logic*, 23:40–50, 2017.

[Montserrat-Adell *et al.*, 2017b] Jordi Montserrat-Adell, Núria Agell, Mónica Sánchez, and Francisco Javier Ruiz. A consensus degree for hesitant fuzzy linguistic decision making. In *Fuzzy Systems (FUZZ-IEEE), 2017 IEEE International Conference on*, pages 1–5. IEEE, 2017.

[Montserrat-Adell *et al.*, 2018] Jordi Montserrat-Adell, Núria Agell, Mónica Sánchez, and Francisco Javier Ruiz. Consensus, dissension and precision in group decision making by means of an algebraic extension of hesitant fuzzy linguistic term sets. *Information Fusion*, 42:1–11, 2018.

[Petty and Gruber, 2011] Jeffrey S Petty and Marc Gruber. in pursuit of the real deal: A longitudinal study of vc decision making. *Journal of Business Venturing*, 26(2):172–188, 2011.

[Rodríguez and Martínez, 2015] R. M. Rodríguez and L. Martínez. A consensus model for group decision making with hesitant fuzzy linguistic information. In *2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, pages 540–545, Nov 2015.

[Rodríguez *et al.*, 2012] R. M. Rodríguez, L. Martínez, and F. Herrera. Hesitant fuzzy linguistic terms sets for decision making. *IEEE Transactions on Fuzzy Systems*, 20(1):109–119, 2012.

[Rodríguez *et al.*, 2013] R. M. Rodríguez, L. Martínez, and F. Herrera. A group decision making model dealing with comparative linguistic expressions based on hesitant fuzzy linguistic term sets. *Information Sciences*, 241:28 – 42, 2013.

[Rodríguez *et al.*, 2014] Rosa M Rodríguez, Hongbin Liu, and Luis Martínez. A fuzzy representation for the semantics of hesitant fuzzy linguistic term sets. In *Foundations of Intelligent Systems*, pages 745–757. Springer, 2014.

[Rodríguez *et al.*, 2016] R. M. Rodríguez, Á. Labella, and L. Martínez. An overview on fuzzy modelling of complex linguistic preferences in decision making. *International Journal of Computational Intelligence Systems*, 9(sup1):81–94, 2016.

[Streletzki and Schulte, 2013] Jan-Georg Streletzki and Reinhard Schulte. Start-up teams and venture capital exit performance in germany: venture capital firms are not selecting on the right criteria. *Journal of Small Business & Entrepreneurship*, 26(6):601–622, 2013.

[Van Osnabrugge, 2000] Mark Van Osnabrugge. A comparison of business angel and venture capitalist investment procedures: an agency theory-based analysis. *Venture Capital: An international journal of entrepreneurial finance*, 2(2):91–109, 2000.

[Zacharakis and Shepherd, 2001] Andrew L Zacharakis and Dean A Shepherd. The nature of information and overconfidence on venture capitalists' decision making. *Journal of Business Venturing*, 16(4):311–332, 2001.