Generation of Visual Representations for Multi-Modal Mathematical Knowledge

Lianlong Wu[1*], Seewon Choi[1*], Daniel Raggi[1*], Aaron Stockdill[2], Grecia Garcia Garcia[2], Fiorenzo Colarusso[2], Peter C.H. Cheng[2], Mateja Jamnik[1]

[1]University of Cambridge, Cambridge, United Kingdom
[2]University of Sussex, Brighton, United Kingdom
{lianlong.wu, daniel.raggi, mateja.jamnik}@cl.cam.ac.uk, sc2179@cantab.ac.uk
{a.a.stockdill, g.garcia-garcia, f.colarusso, p.c.h.cheng}@sussex.ac.uk

Abstract

In this paper we introduce MaRE, a tool designed to generate representations of multiple modalities for a given mathematical problem while ensuring the correctness and interpretability of the transformations between these different representations. The theoretical foundation for this tool is Representational Systems Theory (RST), a mathematical framework for studying the structure and transformations of representations. In MaRE's web front-end user interface, a set of probability equations in Bayesian Notation can be rigorously transformed into Area Diagrams, Contingency Tables, and Probability Trees with just one click, utilising a back-end engine based on RST. A table of cognitive costs, based on the cognitive Representational Interpretive Structure Theory (RIST), that a representation places on a particular profile of user is produced at the same time. MaRE is general and domain independent, applicable to other representations encoded in RST. It may enhance mathematical education and research, facilitating multi-modal knowledge representation and discovery.

Introduction and Theoretical Foundations

In the realm of mathematics, the importance of precise reasoning and well-defined representations, including mathematical graphs, is paramount. In educational settings, it's essential to evaluate the cognitive load these representations impose for assessment and personalised tutoring material development (Larkin and Simon 1987). Similarly, for efficient human-computer collaboration, systems must present information in an easily comprehensible manner and choose representations that enhance communication and learning in technical subjects (Gick 1989; Newell and Herbert Alexander 1972). This necessitates AI systems having an understanding of human reasoning and learning processes (Jamnik and Cheng 2021).

Representational Systems Theory (RST) is a theoretical foundation for studying the structure and transformations of representations (Raggi et al. 2023). As shown in Figure 1, RST can be used for encoding the ways in which humans represent and process information. It is applicable to a variety of formal and informal representation modalities including natural and logical languages, graphs, plots, diagrams and tables. Representational Interpretive Structure Theory (RIST) is our cognitive theory of the interpretive structure of a representation that provides insights into the cognitive costs placed on the human consumer of a particular representation (Cheng et al. 2022; Stockdill et al. 2022).

Related Work AI-generated content systems can produce lifelike images of natural scenes and objects (Rombach et al. 2022). Large language models (LLM) (Brown et al. 2020) with plugins support the generation of tables or tree diagrams, but semantic correctness or logical validity are not guaranteed. Furthermore, these generation processes are implicit and not explainable. However, there is a gap in research focusing on the transformations between formal and diagrammatic notations so that correctness is guaranteed, beyond the production of semantically related images.

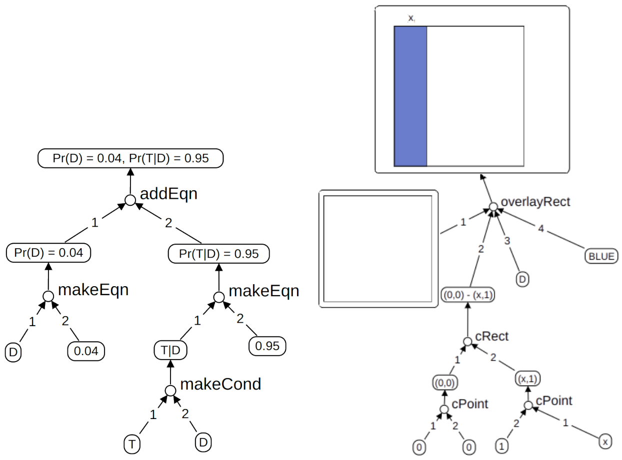
Systems for formal reasoning exist, but are often limited to one formal language or problem type. For example, POTATO (Kovács et al. 2022) is designed for semantic graphs, MM-evocat (Koupil, Bártík, and Holubová 2022) utilises category theory on multi-model database schema, and the Setosa website shows a vivid animation (Victor Powell 2014) for a fixed conditional probability formula only. Wolfram Mathematica (Wolfram Research 2023) necessitates an expert's translation into the Wolfram Language for different diagrams. ChatGPT4 (OpenAI 2023) with the 'Show Me Diagrams' plugin facilitates explanations but cannot ensure generation accuracy.

Our Approach and Highlights We developed a unique diagram generation system called MaRE (Mathematical Representation Engine)[1], where we can input a problem from a variety of domains, in this case a probability problem, in a simple, formal language for conditional probability, and, with just a click, transform from Bayesian notation into three visualised representational systems:

1. Area Diagrams, for representing the probability of events as the relative area of regions;
2. Contingency Tables, for representing joint probabilities in tabular form;
3. Probability Trees, for representing conditional events as branches stemming from a common root.

[*]These authors contributed equally.

[1]The demo video is http://rep2rep.cl.cam.ac.uk/demo/.



(a) Bayesian Notation (b) Area Diagram

Figure 1: Constructions in Representation System Theory.

In MaRE we can input a system of equations for conditional probability in a simple formal language called Bayes. This input gets parsed and encoded as a *construction* – following the RST framework. Our back-end, Oruga (Raggi et al. 2022), takes this construction and automatically transforms it into Area Diagrams, Contingency Tables, and Probability Trees. These transformations are achieved through a process called *structure transfer*, which uses *transfer schemas* to derive the structure of the desired representation. Transfer schemas are expert-developed units of knowledge that capture invariants across different representational systems. Transformations produced by structure transfer are rigorous and transparent. The constructions that result from these transformations are rendered and presented in MaRE. The three alternative representations can be used for elucidating probability theory in educational settings.

Our system has the following key features:

- It can automatically generate a diverse range of diagrams and tables based on the input of probability systems in the form of Bayes expressions.
- It maintains precision and rigour by leveraging RST, while still offering the flexibility commonly associated with the use of diagrams.
- The cognitive analysis based on Representational Interpretive Structure Theory (RIST) provides insights into the cognitive costs of specific representations (Cheng et al. 2021, 2022).

MaRE: Mathematical Representation Engine

We built a novel user interface that can help generate various representations given a mathematical problem.[2] Here we exemplify the domain of probability, where we input a description of the problem in Bayes formal language, and the system can generate various representations with just one click.

User Interface Design The Web user interface[3] of MaRE follows the Single Page Architecture (Emmit A. Scott 2015).

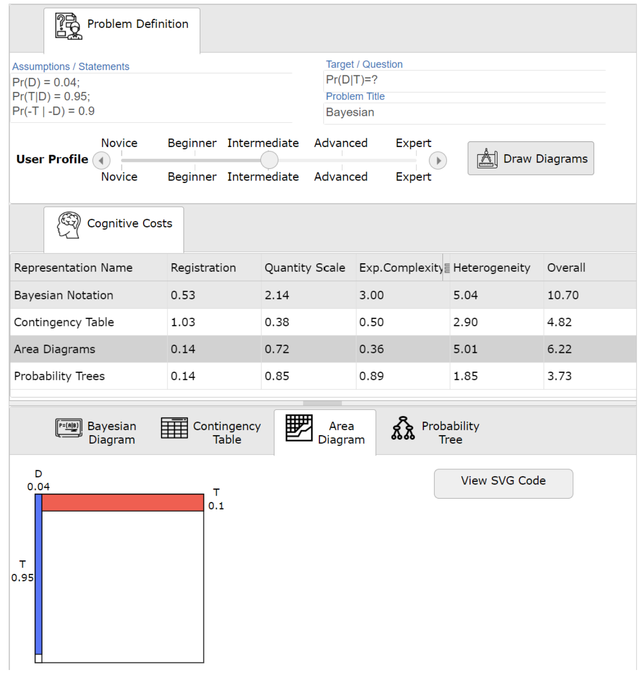


Figure 2: Screenshot of MaRE tool showing Area Diagram.

It is using TypeScript with React as front-end, and NodeJS as API Service, for access to an executable back-end Oruga.

We input a system of probabilistic equations, for example:

```
Pr(D)=0.04; Pr(T|D)=0.95; Pr(-T|-D)=0.9.
```

The string is parsed into a *construction* satisfying the constructor specification of the Bayesian notation.

User Interface Architecture The browser displays the HTML page from a static web server, and the front-end loads dynamic content through the API server. One screenshot is shown in Figure 2. There is a text area that takes a string of our problem description. Below is a user profile slider and one "Draw Diagram" button. The cognitive cost values (computed dynamically for all representations and chosen user profile), and the rendered diagrams are underneath.

There are three forms of output: the HTML snippets are for display in the browser; the PNG pictures can be exported and used for educational material; the SVG source code can be used for further modifications if required.

Generality The RST framework and its implementations are domain-agnostic. So even though MaRE currently only parses and renders representations of probability, other domains and their transformations can be declared over the same back-end. Representations like Set Theory, Venn diagrams, and Arithmetic have been encoded in Oruga.

Discussion and Conclusion

MaRE offers a promising solution to strike a balance between rigour and flexibility in multi-modal transformation. It enables practical applications, demonstrated here, while standing on a theoretically sound and general foundation, to be used in future endeavours in knowledge communication, education and discovery.

[2]https://github.com/lwucam/mare

[3]http://rep2rep.cl.cam.ac.uk/MaRE/

Acknowledgements

This work was supported by the UKRI (EPSRC) Research Grants EP/R030650/1, EP/T019603/1, EP/R030642/1, EP/T019034/1 and G116766.

References

Brown, T. B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; Agarwal, S.; Herbert-Voss, A.; Krueger, G.; Henighan, T.; Child, R.; Ramesh, A.; Ziegler, D. M.; Wu, J.; Winter, C.; Hesse, C.; Chen, M.; Sigler, E.; Litwin, M.; Gray, S.; Chess, B.; Clark, J.; Berner, C.; Mccandlish, S.; Radford, A.; Sutskever, I.; and Amodei, D. 2020. Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33: 1877–1901.

Cheng, P. C.; Garcia, G. G.; Raggi, D.; Stockdill, A.; and Jamnik, M. 2021. Cognitive Properties of Representations: A Framework. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 12909 LNAI, 415–430. Springer International Publishing. ISBN 9783030860615.

Cheng, P. C.-H.; Stockdill, A.; Garcia Garcia, G.; Raggi, D.; and Jamnik, M. 2022. Representational interpretive structure: Theory and notation. In *International Conference on Theory and Application of Diagrams*, 54–69. Springer.

Emmit A. Scott, J. 2015. *SPA Design and Architecture: Understanding single-page web applications*. Simon and Schuster.

Gick, M. L. 1989. Two Functions of Diagrams in Problem Solving by Analogy. *Advances in Psychology*, 58(C): 215–231.

Jamnik, M.; and Cheng, P. 2021. Endowing machines with the expert human ability to select representations: Why and how. *Human-Like Machine Intelligence*, 355–378.

Koupil, P.; Bártík, J.; and Holubová, I. 2022. MM-evocat: A Tool for Modelling and Evolution Management of Multi-Model Data. *International Conference on Information and Knowledge Management, Proceedings*, 4892–4896.

Kovács, Á.; Gémes, K.; Iklódi, E.; and Recski, G. 2022. POTATO: exPlainable infOrmation exTrAcTion framewOrk. In *International Conference on Information and Knowledge Management, Proceedings*, 4897–4901. ACM. ISBN 9781450392365.

Larkin, J.; and Simon, H. A. 1987. Why a Diagram Is (Sometimes) Worth 10000 Words. *Cognitive Science*, 11(1): 65–99.

Newell, A.; and Herbert Alexander, S. 1972. Human problem solving. *NJ: Prentice-hall*, 104(9).

OpenAI. 2023. ChatGPT4, August 3 Version. San Francisco, California, 2023.

Raggi, D.; Stapleton, G.; Jamnik, M.; Stockdill, A.; Garcia, G. G.; and Cheng, P. C. 2022. Oruga: An Avatar of Representational Systems Theory. *CEUR Workshop Proceedings*, 3227: 1–5.

Raggi, D.; Stapleton, G.; Jamnik, M.; Stockdill, A.; Garcia Garcia, G.; and Cheng, P. C.-H. 2023. *Representational Systems Theory: A Unified Approach to Encoding, Analysing and Transforming Representations*. CSLI Press. Forthcoming. Available at arXiv:2206.03172.

Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; and Ommer, B. 2022. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10684–10695.

Stockdill, A.; Garcia Garcia, G.; Cheng, P. C.; Raggi, D.; and Jamnik, M. 2022. Cognitive Modeling of Interpretations of Representations. *The Tenth Advances in Cognitive Systems (ACS) Conference*.

Victor Powell. 2014. Conditional probability - A Visual explanation. Http://setosa.io/conditional/.

Wolfram Research, I. 2023. Mathematica, Version 13.3. Champaign, IL, 2023.