Computational Systems as Co-creative Agents for Visual Humour Generation

José P. Lopes[1,2,*], Pedro Martins[2]

[1]University of Coimbra, Institute for Interdisciplinary Research, Computational Media Design
[2]University of Coimbra, CISUC/LASI – Centre for Informatics and Systems of the University of Coimbra, Department of Informatics Engineering

Abstract

Visual humour in the form of memes transcends geographical and cultural boundaries, enabling individuals to express themselves, share ideas, and participate in online communities. They combine text and visuals, employing humour mechanisms and cultural references to convey messages. The process of meme creation, however, can be complex, especially when trying to convey humour, requiring a combination of creativity, cultural awareness, and technical skills. As a result, the task of creating memes for visual humoristic purposes is not trivial. To address this gap, we leverage generative models to help users ideating and generating visual humour in the format of internet memes by developing two systems that use different interfaces and let the user communicate with a large language model and an image generation model.

Keywords

Meme Generation, Visual Humour, Internet Meme, Co-creativity

1. Introduction

The environment of Web 2.0 gave users the ability to create and share their own memes that can mix pop culture, politics, and participation unpredictably, which shows their versatility in terms of the relationship between different subjects [1]. Despite the growing interest in leveraging generative models for computational internet meme [2, 3], current investigations focus mainly on meme classification and generation, without helping the user ideating their own visual humour concept [4].

2. The systems

To address this gap, we propose two systems using different interaction modalities: one with a conversational interface and another one connecting blocks on a canvas. Both co-creative systems leverage a large language model to help users in visual humour ideation, and an image generation model to materialise their concepts. This is achieved by using the artificial agent to facilitate divergent thinking while relying on the user for convergent thinking to refine and select the most suitable ideas for humorous image generation. The first system, that employs a conversational interface, engages users in a more natural interaction, relying on the popularity of conversational tools, such as ChatGPT and Copilot (Fig. 1). The second system utilises a more visual and spatial interface, with blocks representing specific meme components that are placed on an infinite movable canvas, offering a more visual approach and suggesting content that users can incorporate into their blocks (Fig. 2).

WCDCC 2025: First Workshop on Computational Design and Computer-aided Creativity 2025
*Corresponding author.
joselopes@dei.uc.pt (J. P. Lopes); pjmm@dei.uc.pt (P. Martins)
https://www.cisuc.uc.pt/en/people/joselopes (J. P. Lopes); https://cdv.dei.uc.pt/people/pedro-martins (P. Martins)
0000-0002-8402-6341 (J. P. Lopes); 0000-0002-3630-7034 (P. Martins)

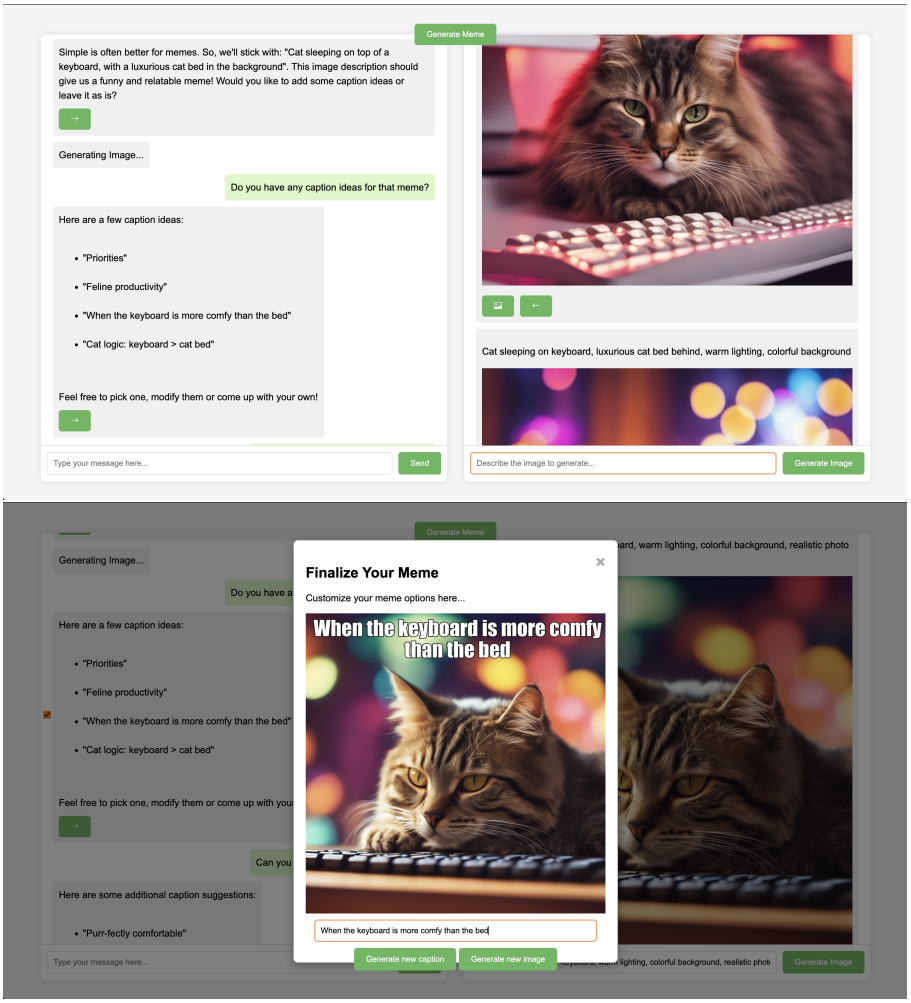

Figure 1: Interface of the Conversational system

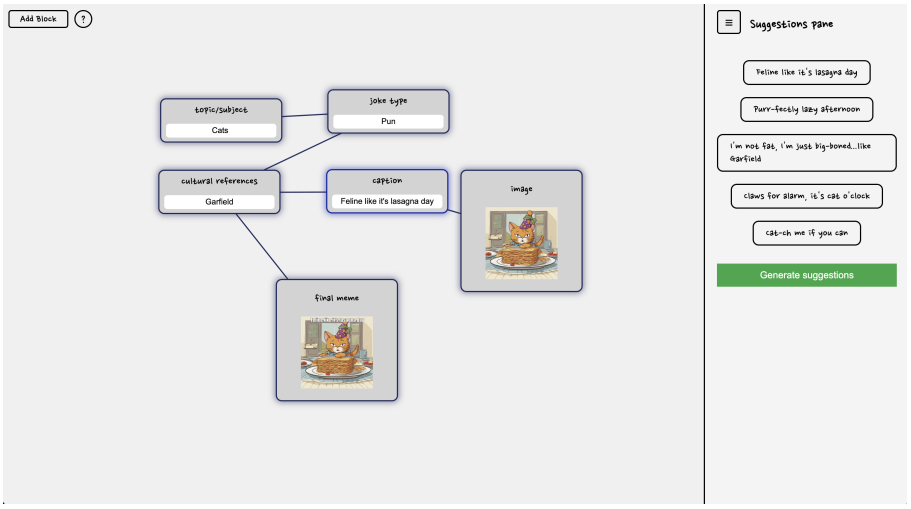

Figure 2: Interface of the Blocks system

2.1. Interaction Flows

To better illustrate how users and the artificial agent collaborate in the two systems, we present the interaction flows. These flows highlight the alternation between user-driven and system-driven actions, and explicitly show iterative loops where ideas can be refined, before the process is finalised.

2.1.1. Conversational System Flow

The interaction in the Conversational System can be summarised as follows:

1. **User provides initial topic or idea.**
2. **System promotes divergent thinking:** proposes humorous text content and questions.
3. **User asks to generate an image.**
4. **System generates an image proposal.**
5. **User validates or refines the output:** evaluates the suggestions and, if unsatisfied, requests new text or images.
6. **System creates a caption and merges it with the image,** finalising the meme once the user accepts the result.

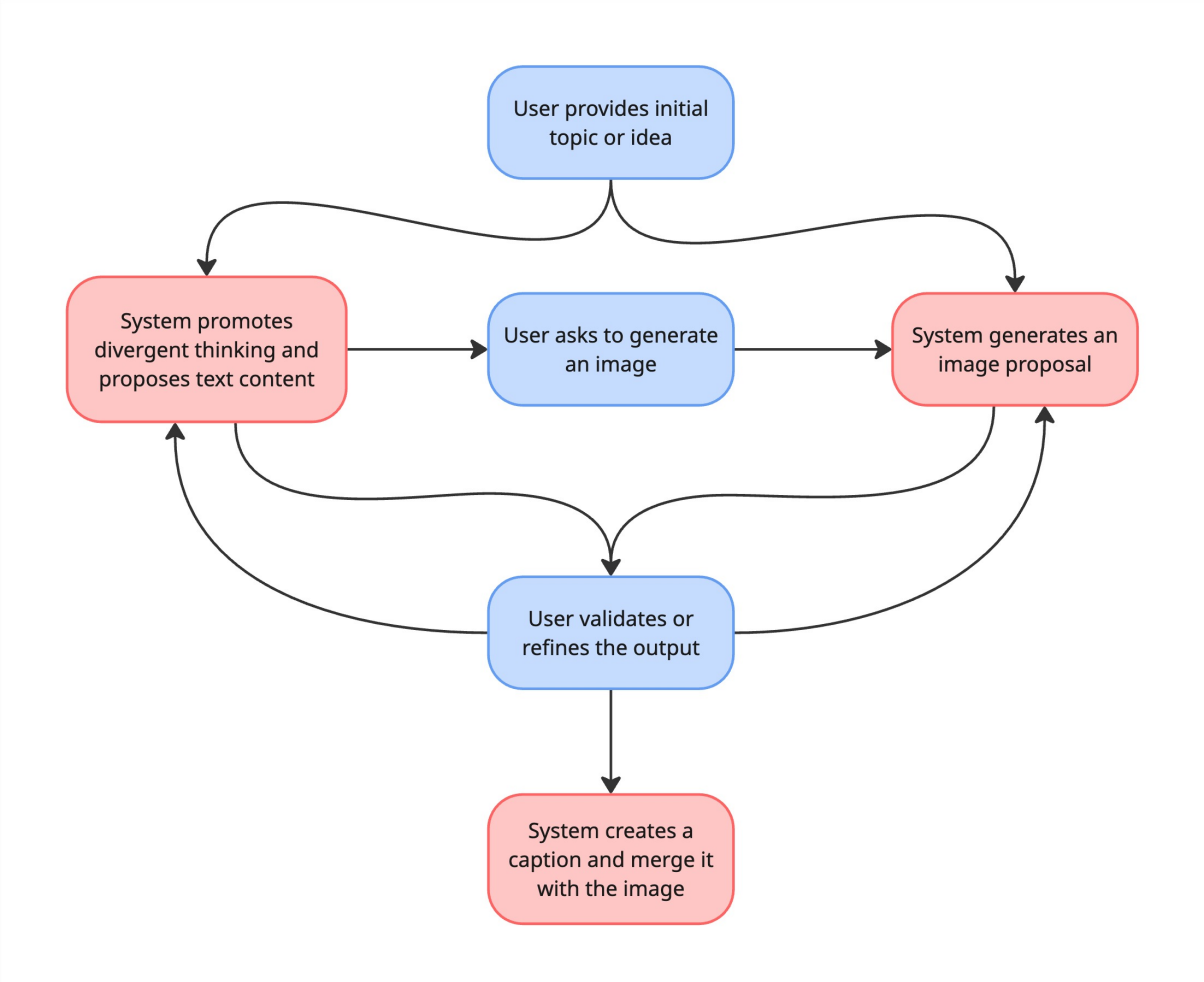

Figure 3: Interaction flow of the Conversational System. Blue nodes = user actions, red nodes = system actions.

2.1.2. Blocks System Flow

The Blocks System interaction emphasises parallel, spatial collaboration. The flow is less linear, but can be summarised as:

1. **User creates a block (textual input).**
2. **System generates suggestions** for that block.
3. **User writes or selects content** for the block.
4. **User creates an image block** and links it with other block types.
5. **System produces outputs** (textual or image) and updates suggestions according to the connected blocks.
6. **User selects content for the block,** refining or combining ideas.
7. **System receives multiple block types** and merges their information.
8. **System merges content from blocks into a final meme,** which the user accepts as the final step.

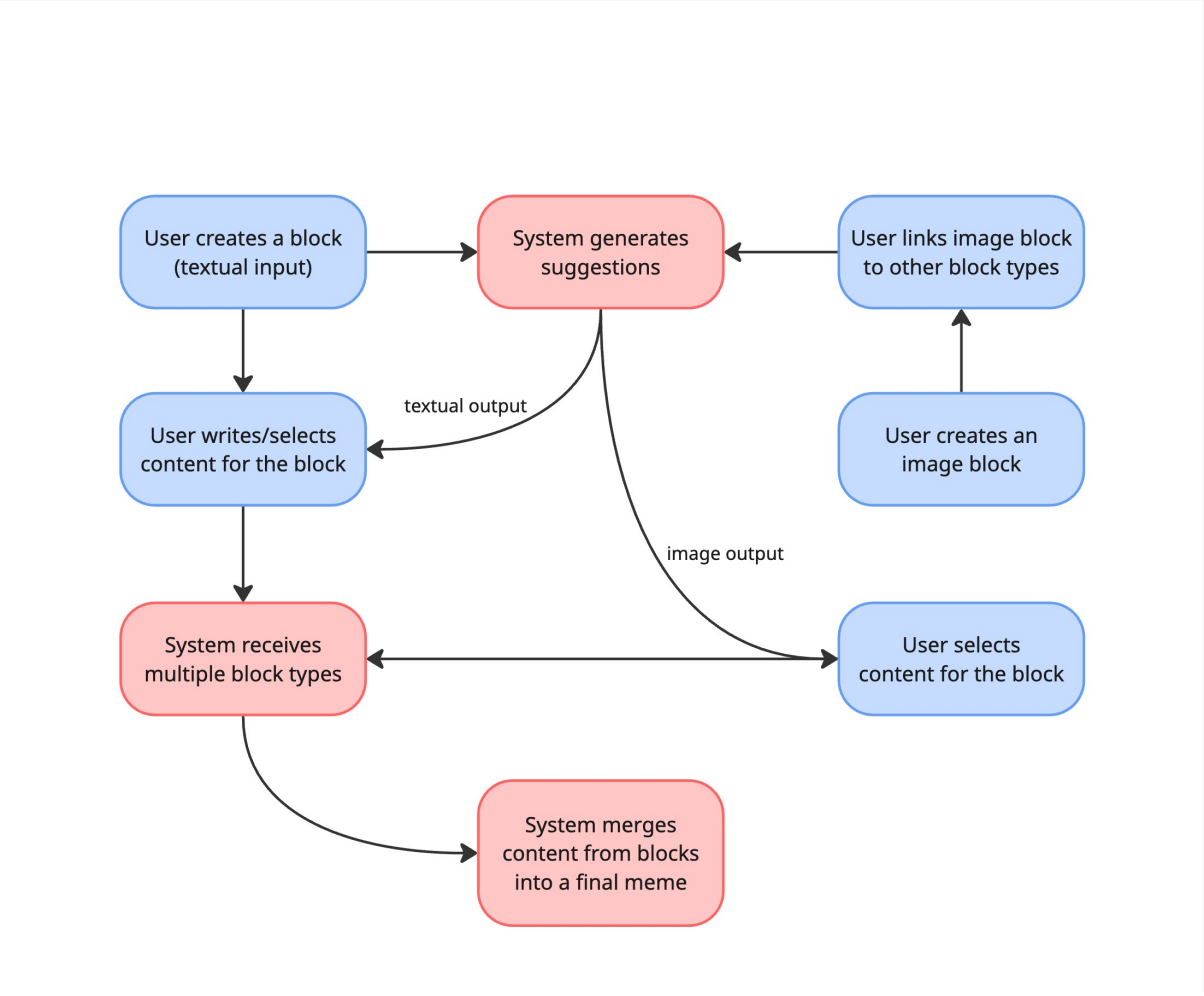

Figure 4: Interaction flow of the Blocks System. Blue nodes = user actions, red nodes = system actions.

Figure 5: Images produced by the Conversational system

Figure 6: Images produced by the Blocks system

3. Results

Results show that the exploratory nature of the Blocks system allowed users to explore more ideas and outcomes, while the natural language interaction of the Conversational System allowed for a more immersive experience. Both systems shown high creativity support scores using the Creativity Support Index [5]. Using a binary evaluation for the outputs generated by users (Figs. 5 and 6), evaluating as funny or not-funny, we were able to circumvent the humour evaluation subjectivity and obtain a humour frequency score [6], which shows that, for both systems, more than half of the outputs are considered funny.

Acknowledgements

This work is funded by national funds through FCT – Foundation for Science and Technology, I.P., within the scope of the research unit UID/00326 - Centre for Informatics and Systems of the University of Coimbra and also supported by the Portuguese Recovery and Resilience Plan (PRR) through project C645008882-00000055, Center for Responsible AI. We would also like to thank João M. Cunha for their valuable guidance and feedback throughout the development of this work.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT and FLUX, in order to: Paraphrase and reword, Improve writing style, Generate images. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] L. Shifman, Memes in Digital Culture, The MIT Press, Cambridge, Massachusetts, 2014.
[2] J. P. Lopes, J. M. Cunha, P. Martins, Stonkinator: An Automatic Generator of Memetic Images, in: Proceedings of the 14th International Conference on Computational Creativity, Canada, 2023.

[3] J. P. Lopes, J. M. Cunha, P. Martins, Computational Creativity in Meme Generation: A Multimodal Approach, in: Proceedings of the 15th International Conference on Computational Creativity, Association for Computational Creativity, Jönköping, Sweden, 2024, pp. 402–406.
[4] R. M. Milner, Pop Polyvocality: Internet Memes, Public Participation, and the Occupy Wall Street Movement, International Journal of Communication 7 (2013) 34.
[5] E. Cherry, C. Latulipe, Quantifying the Creativity Support of Digital Tools through the Creativity Support Index, ACM Transactions on Computer-Human Interaction (TOCHI) 21 (2014) 21:1–21:25.
[6] A. Valitutti, How Many Jokes are Really Funny? Towards a New Approach to the Evaluation of Computational Humour Generators, in: Copenhagen studies in language, 2011, pp. 189–200. ISSN: 0905-7269, Issue: 41.