Mixed-Initiative Co-Creativity

Georgios N. Yannakakis
Institute of Digital Games
University of Malta
Msida, Malta
georgios.yannakakis@um.edu.mt

Antonios Liapis
Center for Computer Games Research
IT University of Copenhagen
Copenhagen, Denmark
anli@itu.dk

Constantine Alexopoulos
Department of Philosophy
University of Edinburgh
Edinburgh, United Kingdom
constantine.alexopoulos@ed.ac.uk

ABSTRACT

Creating and designing with a machine: do we merely create together (co-create) or can a machine truly *foster* our creativity as human creators? When does such co-creation foster the *co-creativity* of both humans and machines? This paper investigates the simultaneous and/or iterative process of human and computational creators in a mixed-initiative fashion within the context of game design and attempts to draw from both theory and praxis towards answering the above questions. For this purpose, we first discuss the strong links between mixed-initiative co-creation and theories of human and computational creativity. We then introduce an assessment methodology of mixed-initiative co-creativity and, as a proof of concept, evaluate Sentient Sketchbook as a co-creation tool for game design. Core findings suggest that tools such as Sentient Sketchbook are not mere game authoring systems or mere enablers of creation but, instead, foster human creativity and realize *mixed-initiative co-creativity.*

Categories and Subject Descriptors

I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems; J.6 [**Computer Applications**]: Computer-Aided Engineering

General Terms

Algorithms, Design, Human Factors

Keywords

mixed-initiative design, computational creativity, lateral thinking, diagrammatic reasoning, co-creation, co-creativity

1. INTRODUCTION

Computer-aided design (CAD) tools have introduced new creation practices through which the computer and the human user collaborate to create new artifacts — be they architectural designs, industrial components, toys or computer games. This paper identifies mixed-initiative co-creation (MI-CC) as the task of creating artifacts via the interaction of a human initiative and a computational initiative. Although mixed-initiative lacks a concrete definition [22], MI-CC in this paper considers both the human and the computer *proactively* making contributions to the problem solution, although the two initiatives do not need to contribute to the same degree. MI-CC thus differs from other forms of co-creation, such as the collaboration of multiple human creators or the interaction between a human and non-proactive computer support tools (e.g. spell-checkers or image editors) or non-computer support tools (e.g. artboards or idea cards).

This paper focuses on game development tasks — in particular on level design — where co-creation occurs between a human and a machine in a mixed-initiative fashion. Level editors such as the *Garden of Eden Creation Kit* (Bethesda 2009) or game engines such as the *Unreal Development Kit* (Epic Games 2009) limit the computer's initiative to interpolations, pathfinding and rendering; while they are very efficient at speeding up game development tasks, human initiative is the sole driver in the creative process. On the other end of the scale, procedural content generators specialized to a type of artifact such as trees with *SpeedTree* (IDV 2002) or First Person Shooter levels with *Oblige* (Apted 2007) can create large amounts of game content but limit the human's initiative to choosing parameters for the generation algorithms; granted that the user has no control during the computer's generative process except before it starts (customizing its parameters) or after it concludes (editing the generated artifact), there is no actual co-creation between human and machine. For the task of game development (and level design in particular), mixed-initiative tools include Tanagra [30] which allows the human designer to specify the position of key platforms in a platformer level with the computational designer generating the remaining level topology, and Sentient Sketchbook [17] which allows human designers to edit a strategy game level while computational creators are simultaneously creating variations of the user's level.

In this paper we argue not solely for the use of mixed-initiative co-creation tools but for their ability to *foster* creativity. We argue that MI-CC can support and realize mixed-initiative *co-creativity*, thus fostering human creativity. Due to the very nature of the MI-CC type examined (level design), creativity in this paper refers to aspects of lateral thinking [10] and diagrammatic reasoning [6]. Our hypothesis is that a human designer interacting with a computational designer that is deemed to be creative is not merely assisted during the creation process; instead, under those circumstances MI-CC *fosters* the designer's creativity. A creative human designer iteratively defines the possibility space of the computational designer which, in turn, influences the *lateral path* of the human designer. While mixed-initiative co-creativity implies that the creativity of both initiatives is fostered, this initial study focuses on and evaluates the human initiative. To support our hypothesis we initially draw from theories on human creativity, outline the theoretical links to MI-CC functionalities and introduce a creativity assessment methodology for MI-CC. The paper's core hypothesis is validated via an empirical evaluation of the Sentient Sketchbook tool [17] for level design, which satisfies the above key condition of mixed-initiative co-creativity. The key findings of the evaluation suggest that MI-CC (via its Sentient Sketchbook instance) can foster aspects of diagrammatic lateral thinking.

2. THE HUMAN INITIATIVE

The understanding of human creativity has relied on diverse philosophical (e.g [33]), neuroscientific (e.g. [9]) and psychological (e.g. [31]) perspectives. While MI-CC can potentially be linked to several theories of human creativity, this paper focuses on aspects of *lateral thinking* and attempts to draw the direct connections between principles of diagrammatic cognitive reasoning and MI-CC.

2.1 MI-CC and Lateral Thinking

Mixed-initiative co-creation — as demonstrated through the level design tool of this paper — is aligned with the general principles of *lateral thinking* [10] and *creative emotive reasoning* [27], the latter being an instance and specialization of the former. Lateral thinking [10] is the process of solving seemingly unsolvable problems or tackling nontrivial tasks through an indirect, non-linear, creative approach. According to De Bono [10], lateral thinking skills can be taught. MI-CC realizes the very nature of lateral thinking which, as a creativity process, is boosted through (increasingly) constrained spaces of solutions [10]. Cocreation with computational creators of visual art and design (e.g. for game level design) encapsulates the very core principles of *diagrammatic reasoning* as human creativity, and especially lateral thinking creativity, is often associated with construction and the principles of customization [10].

The *random stimulus* principle of lateral thinking [2] relies on the introduction of a foreign conceptual element with the purpose of disrupting preconceived notions and habitual patterns of thought, by forcing the user to integrate and/or exploit the foreign element in the creation of an idea or the production of a solution. Randomness within lateral thinking is the main guarantor of foreignness and hence of stimulation of creativity [2]. According to creative emotive reasoning — which enriches the basic notions of lateral thinking with semantic, diagrammatic and emotive dimensions — the creative act is understood as an intervention that results in *re-framing*; frames can be viewed as systems or established routes, that divide the possibility space (e.g. the game design space) into bounded, meaning-bearing sub-areas. On that basis, the random stimulus and the re-framing principles have one element in common: they are enablers of a change in the lateral path. The *re-framing* and the *random stimulus* principles are embedded in the MI-CC paradigm as machine creativity offers heuristically-driven stimuli that are often altered through e.g. mutations within a genetic algorithm; that can, in turn, alter the user's framing on a particular task/problem. An artificial mutation to a visual diagram, an image, or a game map, resembles the random stimulus that can act as a potentiator of creativity and cause an alteration of lateral thinking.

2.2 MI-CC and Diagrammatic Reasoning

Diagrammatic reasoning can be defined as reasoning via the use of visual representations; a cognitive process which is enabled during game level design, interaction design and visual art. These representations can include all forms of imagery incorporating visual features (object shape, size, color, spatial orientation etc.) [6]. Literature suggests that complex information processing is benefited by the use of diagrams, due e.g. to the fact that information in diagrams is indexed by spatial location, thus preserving explicitly the geometric and topological relations of the problem's elements (see e.g. [14]). Diagrammatic reasoning is premised on the background knowledge of the relevant domain, as well as the specific nature of the diagram and its interconnections with the context within which one encounters it [6].

Diagrammatic Lateral Thinking fuses the principles of diagrammatic reasoning and lateral thinking. Diagrammatic lateral thinking builds upon the *extended mind* theory [7] and its core idea is that a diagram, through its use, serves as a vehicle of cognitive processes, embodying the various aspects of the problem. The user's (e.g. designer's) mind is extended onto the diagram and reasoning proceeds through structural (rather than semantic or syntactical) entailment. One therefore *thinks* through the diagram rather than its use as a simple image. According to diagrammatic lateral thinking, the process of constructing a diagram (an image, a map, or a character) is more important that the final product [32]. Moreover, the possibilities one sees for constructing, altering or transforming a given diagram are part of one's comprehension of the diagram itself; the functions of the diagram both on the semantic and pragmatic level are determined in part by these possibilities [28].

MI-CC can not only be viewed as being closely related to lateral thinking but furthermore that it often constitutes a type of diagrammatic lateral thinking: MI-CC occurring through diagrammatic representations (e.g. in level design) offers visual (diagrammatic) alternative paths that satisfy a number of conditions. These define non-linear *lateral paths* within the creative (possibility) space as they promote deep exploration of the space of possibilities which is, in turn, a core lateral thinking characteristic. Diagrammatic lateral thinking, as MI-CC, does not necessarily embed transformational creativity processes as identified by [3]. The MI-CC instance presented in this paper (Sentient Sketchbook) realizes diagrammatic lateral thinking since co-creativity in game level design occurs mainly on the visual (diagrammatic) level, at least in the way levels are presented in Sentient Sketchbook. MI-CC expands the very notion of diagrammatic lateral thinking as it dichotomizes diagrammatic lateral thinking into two main creativity dimensions: one that is based on *analogical* thinking from diagrams and images and one that works purely on the *visual* level through imagistic lateral thinking pathways [27]. In the case of mixed-initiative level design as realized by Sentient Sketchbook [17], MI-CC encapsulates both analogical and visual diagrammatic lateral thinking: the first by constraining the possibility space to playable levels and allowing designers to make analogies to game-specific qualities via diagrams; the latter by targeting visual diversity in the suggestions it provides to the human designer.

3. THE COMPUTATIONAL INITIATIVE

Some of the fundamental questions within computational creativity research are "what does it mean to be creative?" and "does creativity emerge within the individual, the process, the product, or some combination of all three?". The questions are as relevant to human as to machine creativity [3, 8]. Computational creativity, however, seeks creativity generated by, enhanced or fostered via algorithmic means.

The computational creativity literature suggests that *value* (or *usefulness*) and *novelty* are the key elements characterizing a creative process (e.g. see [3]). An autonomous generative system is able to try out exhaustively many possible novel combinations of elements, often resulting in largely uninteresting outcomes or artifacts. For that very reason, computational creativity not only requires the generated artifacts to be *novel*, but also *valuable*. While other aspects of creativity have been discussed and proposed (such as surprise [20]), novelty and value define the common denominators accepted by most theories within computational creativity. If the space of possibilities within MI-CC is constrained for both the machine and the human, the creative process is ultimately of *value* for both given the problem constraints as those are set by either the human or an external observer (e.g. domain expert). Moreover, if the generative process of the machine searches within the constrained space of possibilities for orthogonally possible solutions then the computer interacts with the human user by offering both *useful* and *novel* suggestions throughout the creative process [3]. The end outcome of MI-CC (both novel and useful) is ultimately a result of iterative co-creation. The autonomous creative system, in that case, finds novel ways to navigate a search space, by e.g. looking at orthogonal aspects of the human creative process, which are suggested back to the human.

Computational creativity has been classified by Boden [3] in three types: *combinatorial*, *exploratory* and *transformational*. Combinatorial creativity revolves around the combination of different elements which is often trivially accomplished by a computer. Computers are also well suited for exploratory creativity, which involves traversing a well-defined search space. In contrast, transformational creativity requires the computer to 'break the rules' of that pre-existing conceptual space. Among the three types of computational creativity identified by Boden, MI-CC realizes mainly *exploratory* creativity. While it could potentially achieve transformational creativity, mere exploration of the solution space can often result in more creative outcomes than transformation [4, 24]. Pease et al. [23] provide the example of an unusual but legal chess move as often being more creative than changing the rules of chess.

According to Bundy [4] an outcome is considered creative if the possibility space in which it lies is large (and complex) and if it is generated from a little explored area. MI-CC tools that generate feasible solutions within a small feasible space capture the complexity expressed by Bundy. The harder it is to find a solution within a complex and small feasible search space, the more novel it is deemed [4]. The notion of complexity has also been expressed via a number of alternative computational metrics including rarity and impressiveness [15] that can be considered in a MI-CC tool which involves diagrammatic aspects of creativity.

Autonomous creative systems in the form of procedural content generation (PCG) have been used by the game industry in specialized roles to create engaging but unpredictable game experiences, and to lessen the burden of manual game content generation through automating parts of it [34]. PCG is the backbone and core technique of several MI-CC tools for game design [29, 30, 17]. From a computational creativity perspective, while PCG can be viewed as artifact generation, PCG algorithms are rarely classified as creative. Evaluating commonly used PCG algorithms based on the creative tripod of skill, appreciation and imagination [8], a case can be made that most existing algorithms possess only skill. This paper attempts to evaluate the degree to which a PCG algorithm which generates valuable and novel content for the human designer to consider can contribute to human creativity. The next section discusses approaches for evaluating the creativity of the MI-CC paradigm, investigating both initiatives (human and computer) and their fusion.

4. HOW TO EVALUATE THE CREATIVITY OF MI-CC

Evaluating the impact of a mixed-initiative tool on computational creativity or human creativity is far from trivial and needs to be further supported empirically, via quantitative metrics and qualitative studies. Proposals have been put forth for evaluating the results of an autonomous computational creator [26] as well as the usability of "traditional" CAD tools [5, 11]. While any of these proposed metrics could arguably be relevant when evaluating MI-CC tools, they fail to capture the impact of the proactive computational initiative of MI-CC on human creativity and vice versa — computational creativity assumes minimal human initiative and human-computer interaction assumes minimal computational initiative. This paper offers a first step towards the evaluation of mixed-initiative co-creation, focusing on its ability to foster human creativity. While the interaction between human and computer arguably also fosters the computational creativity of the MI-CC tool, supporting or evaluating such a claim is outside the scope of this paper as it requires different theoretical frameworks and experiments.

Generally there are two types of creativity evaluation to be considered when a computational creator is involved such as in MI-CC: the evaluation of the final (or possibly intermediate) outcomes and the evaluation of the co-creative process for the generation of outcomes, solutions, or items. The former can be evaluated through a number of heuristics for the task at hand (such as novelty and usefulness) or through crowdsourced estimates of creativity from a human (or even a computational) audience [26]. The latter is less straightforward to evaluate as the exact human creativity processes are either completely unknown or only partially known [12]. It would, thus, require either (a) identifying milestones within the (co-)creative process through the heuristics of novelty and value (and surprise [20] or other relevant heuristics) which may approximate aspects of the underlying creative process or (b) some type of meta-level (or self-predictive) mechanism of the quality of the process or, alternatively, (c) an evaluation based on a temporal model of the co-creative process [21]. We argue that, supplementary to fostering the creativity of the final co-creative outcome, MI-CC supports and fosters the *creative process* towards that outcome.

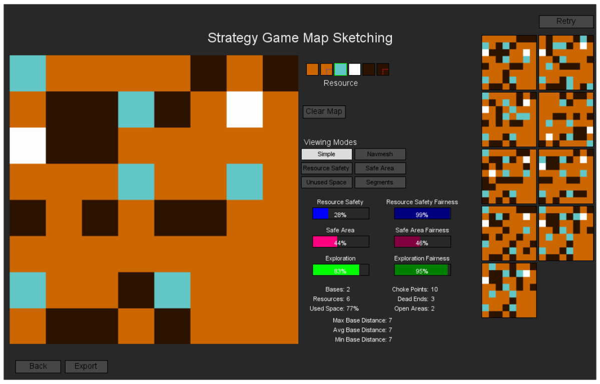
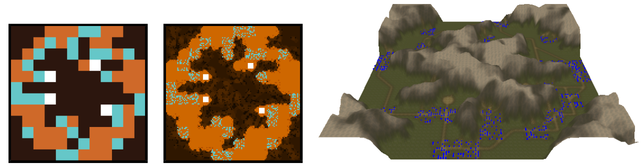


Figure 1: The interface of Sentient Sketchbook while a user edits a map sketch. The editing window covers the left half of the screen, while the computer generated suggestions are shown in the far right edge. Between the two are the tile palette, options for alternative displays of the map, and evaluations of the user-created map.



(a) Sketch of a strategy game level. (b) Final game level of Fig. 2a. (c) 3D strategy game level of Fig. 2a.

Figure 2: Sample map sketch of a strategy game level and the maps it creates. Tiles can be impassable (dark) passable (orange), bases (white) or resources (cyan).

Within MI-CC, it is the human creator herself that implicitly judges the quality and degree of use of the computational creator. However, it is an unbiased (and potentially naïve) human audience that judges the quality of this use within the creative lateral path. On that basis the creativity capacity of MI-CC is evaluated in two quantitative ways in this study: through (a) the *degree of use* of the computational creations and (b) the *quality of their use* within the lateral path of creation. For this purpose, a human audience (game designers) is used both to reveal the usefulness of MI-CC (through its use) and to annotate *milestones* (i.e. enablers of a change) in the lateral path that have potentially been caused by the interaction with the computational creator. This paper investigates and evaluates the Sentient Sketchbook tool for game level design as an instance of MI-CC.

5. SENTIENT SKETCHBOOK

Sentient Sketchbook is a mixed-initiative tool for the design of game levels [17]. Using map sketches as a low-fidelity representation of a game level, Sentient Sketchbook allows its users to create content in substantially less time than if they were given full control of the design process. Map sketches are minimal abstractions of game levels consisting of a small number of tiles, and can represent several types of game levels [19]; this paper will focus on levels for strategy games. A map sketch has tiles that allow movement (passable) or obstruct it (impassable), and passable tiles can contain player bases and resources. These simple sketches are quick to comprehend, evaluate, and edit, allowing the user to playfully interact with the tool and try different level setups with minimal investment in time and effort. Once the sketching process is complete or at any time during the interaction, the computer can convert these simple map sketches into complete, playable levels in 2D or 3D (see Fig. 2).

Apart from the conversion of sketches into playable game levels, the computer assists the human designer via an intuitive interface and automated evaluation of their sketches on domain-specific properties. Due to the small size and complexity of map sketches, evaluating navigational and topological properties of the user's level is lightweight and can be performed on real-time with every user interaction. The evaluations are generic and pertain to proximity, branching paths and balance [19]. Apart from numerical evaluations, which are shown as colored horizontal bars ranging from 0% to 100% (see Fig. 1), the tool also offers visual representations, including a navigational grid, optimal paths between bases, or "safe" resources around each player base.

The most promising feature of Sentient Sketchbook for prompting actual human-machine co-creativity is the generation of suggestions by the computer. These suggestions are generated in real-time while the users edit their levels; up to twelve suggestions are presented on the edge of the user's editing window (see Fig. 1), and the user can select any of these suggestions to substitute their current map sketch. Suggestions are generated via genetic algorithms, a computational simulation of Darwinian evolution centered around *survival of the fittest*. In Sentient Sketchbook, map sketches are evolved via a feasible-infeasible two-population genetic algorithm [13] which ensures that all shown suggestions are constrained to be playable, i.e. that all special tiles are connected via a passable path. The genetic algorithm evolves permutations of the user's sketch either to optimize the domain-specific level evaluations used by the tool, or to create maps that are visually diverse from the user's sketch. Due to the need for immediate user feedback, the evolutionary sprint is short; this results in maps which may not be optimal, yet retain sufficient structural properties of the original user's sketch to appear "familiar". In the current version of the tool, 6 of the generated suggestions evolve to optimize numerically defined fitness functions which match the 6 evaluations of level quality shown to the user via the interface. The remaining 6 generated suggestions evolve towards visual diversity via feasible-infeasible novelty search [16] which creates diverse yet feasible individuals. While these suggestions do not explicitly improve a game level as they have no notion of level quality — except for the constraint on playability — they have been shown to provide inspiration in cases where a user does not have a level concept in mind.

6. EVALUATION

For an initial study of Sentient Sketchbook, five expert users (employed in the game industry) were asked to create strategy game levels using the tool [17]. These expert users were acquaintances of the authors who were sent the tool via e-mail with a request to create a few maps and provide feedback; the participants' usage logs were sent back via e-mail after approximately a week, during which time they used the tool at their leisure. Additional questions on usability were asked and answered via e-mails.

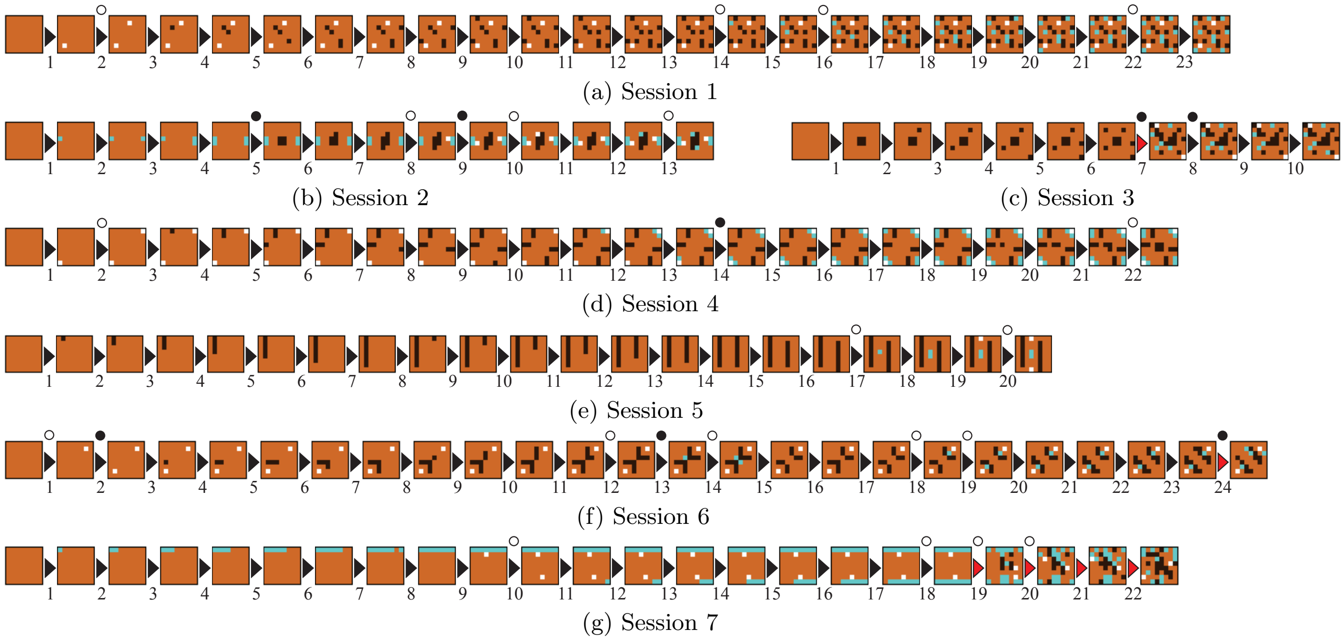


Figure 3: Complete creation paths of seven design sessions on small maps, displaying all steps taken. Each step corresponds to a user's interaction, and causes the transition from the previous map to the next. Steps where a generated suggestion replaced the user's sketch are shown as red arrows. Steps tagged as milestones in the creation process by two designers are indicated with a white circle above the arrow; those tagged by three designers are indicated with a black circle.

Overall, participants of this user survey responded positively to the editing interface, while the evaluations and different visualizations of map properties were considered relevant to the domain of strategy games and helpful to the creative process. This section follows the creativity assessment scheme introduced in Section 4 and presents findings on the use and the quality of use of the computational initiative of Sentient Sketchbook.

6.1 Degree of Use

Regarding the suggestions generated by the computational creator, participants made use of them in 15 out of 24 design sessions. According to user feedback, cases where suggestions did not prove useful included (a) instances where a user had preplanned a map layout before starting their design session, (b) instances where the strict visual symmetries of human maps were not retained in the generated suggestions, (c) instances where suggestions did not retain the number of bases of the user's sketch and (d) instances where suggestions appeared disconnected to what the user was focusing on at that time. While there is a clear way of amending instances of type (c) and — to an extent — (b), instances of type (a) will never benefit from MI-CC processes, while instances of type (d) require an elaborate model of the designer's goals, processes and preferences [18].

In the design sessions where computer-generated suggestions were selected to substitute the user's sketch, participants predominantly used map suggestions evolved to maximize a map property. This verifies, in a way, that the tool's map evaluations are appropriate to the domain of strategy game level design and useful to designers. Suggestions evolving to maximize a map property were selected mostly towards the end of the level design process, in order to "finetune" an almost finished map sketch. From the current findings, it is hard to deduce whether this behavior is an inherent property of the creative process itself, i.e. that users focused on the task or vision they had contrived and only noticed evaluations or suggestions once their own creative drive had run its course, or whether it is an artifact of the user interface and quality of suggestions. While map suggestions evolving towards visual diversity were selected only in 6 out of 22 cases where suggestions were used, such suggestions were predominantly selected during the early stages of the creative process. This deviation indicates that while users can use targeted suggestions, which are conscious of strategy game level quality, to "finetune" a map created predominantly by their own creative spark, suggestions which target visual novelty, despite perhaps being worse in terms of level quality than those explicitly targeting it, can provide inspiration and alter the visual lateral path of the designer.

6.2 Quality of Use

To illustrate the impact of mixed-initiative design in the creative process — and due to space considerations — we will only focus on small maps of 64 tiles (8 by 8), as they are more concise in terms of discrete user actions taken from empty map to final map. During the user study detailed both above and in [17], 7 design sessions on small maps were undertaken by multiple expert designers.

6.2.1 *Qualitative Observation of Creation Paths*

The creation path of each map is shown on Fig. 3; every map instance, before and after the user's action, is displayed sequentially from the session's start to its finish. With the exception of session 1, the patterns in user actions in the sessions shown indicate a preference towards symmetry, not only on the final map but also on the process; that is shown often in the initial placement of bases (sessions 4 and 6), re-

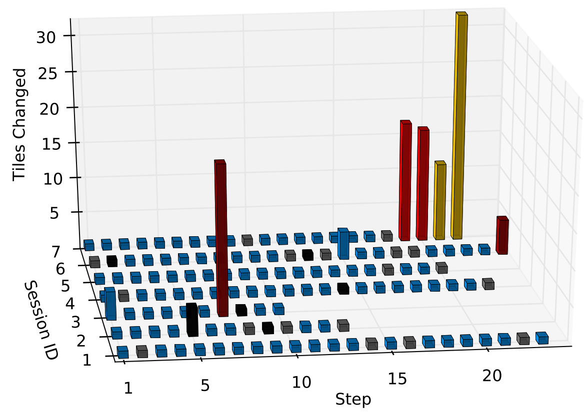


Figure 4: Tiles changed on each step for the design sessions shown in Figure 3. Steps tagged as milestones by 2 users are shown in gray, and by 3 users in black. Steps with suggestions are shown in yellow (not tagged by multiple users), red (tagged by 2 users) and dark red (tagged by 3 users).

sources (session 2) or impassable regions (sessions 3 and 5). In sessions 1, 2, 4 and 5, no computer-generated suggestions are chosen to replace the user's design; the designer's tendency towards symmetry is never broken (excluding session 1, which did not include symmetry at any point in the process). On the other hand, in sessions 3 and 6 the designer used a suggestion to change their current design once (at steps 7 and 24 respectively), while in session 7 the designer used the suggestions multiple times in a row (steps 19 to 22), changing their design thoroughly. In session 6, the suggestion caused a minor change in the map structure: while the diagonal symmetry of the map is broken, the suggestion does not seem to result in a better map or inspire the user to continue working on it. In session 3, the suggestion caused all necessary game elements to be added to the map (i.e. two bases and multiple resources), acting as a form of shortcut to speed up the creation process; the designer then proceeds to manually "correct" the generated suggestion by creating multiple paths between bases and balancing the placement of resources around each base. Finally, session 7 is interesting in the fact that the user created a very symmetrical initial map and then proceeded to explore other possibilities via the suggestions; steps 19 and 22 saw the use of suggestions targeting novelty, which lead to big changes in the map, while steps 20 and 21 saw the use of suggestions targeting specific map qualities, which caused more subtle changes which can be considered "finetuning".

6.2.2 *Quantitative Evaluation of Creation Paths*

While the qualitative evaluation of the design process is quite informative, the question of quantitatively evaluating the creation path (and its creativity) remains to be answered. A mathematically defined heuristic could potentially be used to evaluate the process; in the domain of tile-based map sketches, the number of tiles changing within a step could perhaps be a good abstraction of how humans qualitatively measure map change. Fig. 4 displays the tiles changed from one step (i.e. one user action). As the interface of Sentient Sketchbook mostly supports "painting" a single tile, most user actions result in one tile changed; all steps of sessions 1, 4 and 5 see one tile changed per step. The designer can also paint a larger passable or impassable area, leading to a handful of tiles changing (step 5 of session 2, step 1 of session 3 and step 15 of session 6). However, only the replacement of the user's sketch with a computer-generated suggestion causes large changes in the number of tiles. For instance, 21 tiles change at step 7 of session 3, when the user chooses a computer-generated suggestion that adds bases and resources; despite the fact that the suggestion was not evolved towards visual diversity, the large change is due to the playability constraint that computer-generated maps must possess multiple bases and resources. In session 6, the generated suggestion at step 24 does not change many of the map's tiles, as is directly visible in Fig. 3. In session 7, generated suggestions evolved via novelty search change the map substantially (steps 19 and 22) while those generated towards optimizing a map quality less so (steps 20 and 21). The fact that novelty search makes more significant changes to maps is not surprising, since it explicitly targets tile difference in the novelty score of the genetic algorithm.

6.2.3 *Human Audience*

Another evaluation which combines both the subjective properties of qualitative human evaluation and quantitative, data-driven evaluations comes from the feedback of a human audience regarding the creation process. In this context, the expert users participating in the presented user survey were asked to tag which of the steps displayed in Fig. 3 constituted *milestones* in the creation process — i.e. enablers of lateral path change. All survey participants were asked to tag all creation paths (excluding theirs) for small maps. The user's responses were unsurprisingly varied: some designers tagged numerous steps as milestones (as high as 8.4 milestones on average per session) while others identified milestones sparsely (1.17 milestones on average per session). To the same effect, different designers arguably selected milestones based on different underlying criteria. This study will investigate steps classified as milestones by two or more users, excluding the map's creator. Both Fig. 3 and Fig. 4 highlight which steps were classified as milestones by two or three designers (no milestones had the consensus of all four designers). While many of these milestones are commonsensical to a casual, human observer, it is interesting to note that they are not often detected by the tile change heuristic. The steps which place the first or second base (e.g. step 2 of session 4, step 10 of session 7) or the steps which turn an asymmetrical map into a symmetrical one (e.g. step 14 of session 4, step 20 of session 5) are often identified as milestones, yet they involve a single tile change and are not identified as significant by the heuristic. However, the designer's analogical reasoning behind the pure visual representation of the strategy game level identifies that bases are significant for gameplay and that symmetrical maps result in "fair" games. Designers are thus not tagging milestones solely for their potential as visual diagrammatic lateral paths, but also as analogical diagrammatic lateral paths. Although several steps were tagged as milestones by multiple users, suggestions are particularly prevalent among them, considering their rarity. Out of the 6 computer suggestions in the sessions investigated, 4 were tagged as milestones by two or more users, out of a total of 28 milestones on 134 steps, while 2 suggestions were tagged as milestones by three users out of a total of 8 milestones on 134 steps.

It is important to note that identifying a computer generated suggestion is not particularly difficult for casual observers and users of Sentient Sketchbook alike; the tile change heuristic can also detect suggestions with a low margin of error. The fact that generated maps are not symmetrical is a tell-tale sign of a computer generated suggestion, as well as the fact that multiple tiles change in a single step. While in cases where designers were focusing on symmetrical maps these generated asymmetrical patterns did not always resonate with them, in the context of MI-CC the fact that generated suggestions do not appear similar to something a human designer would have created is not in any way detrimental. Instead, the asymmetries act as a stimulus necessary to disrupt the user's habitual pattern of thought and, to a degree, the ingrained attraction of humans to symmetry [1, 25]. By targeting map qualities such as balance, generated suggestions which are asymmetrical, yet of high quality and balanced, disrupt preconceived notions of level designers that "fairness" between opponents in a strategy game can only be achieved through symmetry in the levels. As can be expected, the foreignness of the generated suggestions is often rebelled against and suggestions are ignored, especially in cases where users are not necessarily open to non-linear diagrammatic lateral thinking. However, even the presentation of such suggestions while the user is editing the level can lead to a certain didactic experience in diagrammatic lateral thinking skills; as a participant of the user survey commented regarding their experience using Sentient Sketchbook, "I've started to like the tool a lot. Instead of doing something symmetrical and intricate, I started just working with some high level abstract ideas for the maps, like 'this one should be scattered' or 'in this one the players should be 3 vs. 1' or whatever. Suddenly the tool starts pushing you to places you didn't really consider."

7. DISCUSSION AND CRITIQUE

The main line of criticism that can be leveled against the notion of mixed-initiative co-creativity can be summed up in the following objection: "The user merely *chooses* from alternatives, i.e. the suggestions presented by Sentient Sketchbook. There is no *real* creativity on the part of the human user." An additional key critique is: "The computer merely follows a random process". While we do not focus on or evaluate the computational end of MI-CC in this paper (latter objection), we will try to decompose and answer to the first objection by analyzing its different parts.

The theoretical foundation of diagrammatic lateral thinking, and consequently of MI-CC, is the notion of the *extended mind* [7]. An external object (e.g. a notebook, software, etc.) that is consistently involved in and relied upon in order for a human subject to perform some cognitive or reasoning function, can be understood as an essential constituent of that subject's mind. If we take the example of a person suffering from memory problems who consistently uses and relies upon a notebook in order to carry out everyday tasks (either simple or complex), then we can understand the example's notebook as a constituent of that person's mind, an *extension* of it. It is in this sense that the MI-CC tool presented in this paper, which aims for more use, and is meant to be consistently relied upon in a creative process (which involves cognitive/reasoning skills), defines such an *extension.* This is important as it considerably reduces the distance between human user and machine, which the objection relies upon. There is no outside party doing the real work, and a human subject "merely choosing," but an integrated cognitive/reasoning system.

This distance is further reduced when we take into account the iterative process of suggestion generation. As the creative process unfolds, the user is constantly guiding the MI-CC tool as it, in turn, guides the human user. The notion of co-creation is key. Furthermore, this casts doubts as to whether the simple concept of *choice* is applicable here. The suggestions (and consequent selections) are a product of a human-machine interaction, and they have meaning as links within a chain, that takes us from the user's initial designs/intentions to the final creative product. The human user is not "merely choosing" but shaping the space which generates the suggestions themselves. Even if we disregarded this and use choice in the limited interpretation of this objection, we still find a place for "merely choosing" in our common-sensical notion of a creative process. A significant part of any creative process consists of choosing between alternatives. What makes it creative is how we further exploit this choice and the reasons for choosing as we chose.

Even if we did not view MI-CC under the extended mind theory, we can still understand the human user and the computational creator as forming a group. Within a group, co-creativity can take many forms and is not constrained by the narrow, individualistic perspective underlying the objection under discussion. In a group, one is often labeled creative if she stirs the group's creativity by making the right choices between alternatives produced by the group as a whole.

One obvious critique against the evaluation scheme used is the lack of a control experiment. While controlling for the use of the computational initiative in a very direct manner could potentially provide additional information on the usefulness of MI-CC, it would not necessarily address the question of fostering creativity. The proposed experimental protocol instead allows users to freely choose to apply or ignore suggestions provided to their designs and then, in a post-experience manner, evaluate the milestones of other designer's lateral paths. Such an evaluation validates the key hypothesis of this paper directly.

Another critique that MI-CC evaluation faces is its very focus on the creative process — and the use of the computational creator — instead of the evaluation on the final artifact [26]. While the outcome of a creative process may reveal aspects of creativity, we argue (and literature suggests [23]) that it is the lateral path itself (i.e. the creative design process) that is more important and potentially more relevant to evaluate. The iterative co-creation properties of MI-CC for creative design (such as level design and visual art) often allow for transparent and observable creation paths to be available and, thus, offer richer underlying information about the co-creativity of humans and machines.

8. CONCLUSIONS

In this paper we attempt to shed light on the creativity capacity of mixed-initiative co-creation in and for games. We thus situate mixed-initiative co-creation within the literature of human and computational creativity, identifying clear links to both diagrammatic lateral thinking and exploratory creativity. As MI-CC uniquely fuses aspects of human and computational creativity, we propose an assessment methodology of the potential of MI-CC as an enabler for diagrammatic lateral path change. The evaluation is performed on the Sentient Sketchbook MI-CC tool and the key results indicate that the computational initiative is useful to humans on the task of strategy game level design and, most importantly, that the selected computer-generated level designs define milestones in the creative lateral path of the human designers surveyed.

Through the study of mixed-initiative co-creativity we aim to advance our understanding of human creativity and develop software capable of fostering it. The main motivation behind MI-CC is that the potential of human creativity is often undermined because users (e.g. designers or players) lack the appropriate co-creation tools. The core findings of this paper suggest that the mixed-initiative co-creation paradigm is more than an *enabler* for human creativity, a mere computer-assisted design tool or a facilitator of human co-creativity such as any CAD tool or an implicit co-creation enabler such as social media. Instead, through the mixed-initiative perspective we assume an autonomous computational system that explores the possibility space in its own ways as guided by human lateral decisions during the creative process, realizing and fostering human-machine *co-creativity*.

9. ACKNOWLEDGMENTS

We would like to thank the participants of the user study for their valuable feedback. The research is supported, in part, by the FP7 ICT project C2Learn (project no: 318480) and the FP7 Marie Curie CIG project AutoGameDesign (project no: 630665).

10. REFERENCES

[1] R. Arnheim. *Art and visual perception: a psychology of the creative eye*. University of California Press, revised and expanded edition, 2004.

[2] M. Beaney. *Imagination and creativity*. Open University Milton Keynes, UK, 2005.

[3] M. A. Boden. *The creative mind: Myths and mechanisms*. Routledge, 2003.

[4] A. Bundy. What is the difference between real creativity and mere novelty? *Behavioral and Brain Sciences*, 17(03):533–534, 1994.

[5] E. A. Carroll, C. Latulipe, R. Fung, and M. Terry. Creativity factor evaluation: Towards a standardized survey metric for creativity support. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition*, pages 127–136, 2009.

[6] P. C.-H. Cheng, R. K. Lowe, and M. Scaife. Cognitive science approaches to understanding diagrammatic representations. *Artificial Intelligence Review*, 15(1-2):79–94, 2001.

[7] A. Clark. *Being there: Putting brain, body, and world together again*. MIT press, 1998.

[8] S. Colton. Creativity versus the perception of creativity in computational systems. In *AAAI Spring Symposium: Creative Intelligent Systems*, pages 14–20, 2008.

[9] A. R. Damasio. Some notes on brain, imagination and creativity. *The origins of creativity*, pages 59–68, 2001.

[10] E. De Bono. *Lateral thinking: Creativity step by step*. Harper Collins, 2010.

[11] T. Hewett, M. Czerwinski, M. Terry, J. Nunamaker, L. Candy, B. Kules, and E. Sylvan. Creativity support tool evaluation methods and metrics. In *Report of Workshop on Creativity Support Tools*, pages 10–24, 2005.

[12] D. R. Hofstadter. *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought*. Basic Books, 2008.

[13] S. O. Kimbrough, G. J. Koehler, M. Lu, and D. H. Wood. On a feasible-infeasible two-population (fi-2pop) genetic algorithm for constrained optimization: Distance tracing and no free lunch. *European Journal of Operational Research*, 190(2):310–327, 2008.

[14] J. H. Larkin and H. A. Simon. Why a diagram is (sometimes) worth ten thousand words. *Cognitive science*, 11(1):65–100, 1987.

[15] J. Lehman and K. O. Stanley. Beyond open-endedness: Quantifying impressiveness. In *Artificial Life*, volume 13, pages 75–82, 2012.

[16] A. Liapis, G. Yannakakis, and J. Togelius. Enhancements to constrained novelty search: Two-population novelty search for generating game content. In *Proceedings of Genetic and Evolutionary Computation Conference*, 2013.

[17] A. Liapis, G. Yannakakis, and J. Togelius. Sentient sketchbook: Computer-aided game level authoring. In *Proceedings of ACM Conference on Foundations of Digital Games*, 2013.

[18] A. Liapis, G. N. Yannakakis, and J. Togelius. Designer modeling for personalized game content creation tools. In *Proceedings of the AIIDE Workshop on Artificial Intelligence & Game Aesthetics*, 2013.

[19] A. Liapis, G. N. Yannakakis, and J. Togelius. Towards a generic method of evaluating game levels. In *Proceedings of the AAAI Artificial Intelligence for Interactive Digital Entertainment Conference*, 2013.

[20] L. Macedo and A. Cardoso. Modeling forms of surprise in an artificial agent. *Structure*, 1(C2):C3, 2001.

[21] M. L. Maher, K. Brady, and D. H. Fisher. Computational models of surprise in evaluating creative design. In *Proceedings of the Fourth International Conference on Computational Creativity*, 2013.

[22] D. Novick and S. Sutton. What is mixed-initiative interaction? In *Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction*, 1997.

[23] A. Pease, D. Winterstein, and S. Colton. Evaluating machine creativity. In *Workshop on Creative Systems, 4th International Conference on Case Based Reasoning*, pages 129–137, 2001.

[24] J. Pind. Computational creativity: What place for literature? *Behavioral and Brain Sciences*, 17(03):547–548, 1994.

[25] V. S. Ramachandran and W. Hirstein. The science of art: a neurological theory of aesthetic experience. *Journal of consciousness Studies*, 6:15–51, 1999.

[26] G. Ritchie. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1):67–99, 2007.

[27] T. Scaltsas and C. Alexopoulos. Creating creativity through emotive thinking. In *Proceedings of the World Congress of Philosophy*, 2013.

[28] A. Sloman. *Diagrams in the Mind?* Springer, 2002.

[29] R. M. Smelik, T. Tutenel, K. J. de Kraker, and R. Bidarra. A declarative approach to procedural modeling of virtual worlds. *Computers & Graphics*, 35(2):352–363, 2011.

[30] G. Smith, J. Whitehead, and M. Mateas. Tanagra: Reactive planning and constraint solving for mixed-initiative level design. *IEEE Transactions on Computational Intelligence and AI in Games*, (99), 2011.

[31] R. J. Sternberg. *Handbook of creativity*. Cambridge University Press, 1999.

[32] A. Vile and S. Polovina. Thinking of or thinking through diagrams? The case of conceptual graphs. In *Thinking with Diagrams Conference, The University of Wales, Aberystwyth*, 1998.

[33] L. Wittgenstein. *Philosophical investigations*. John Wiley & Sons, 2010.

[34] G. N. Yannakakis and J. Togelius. Experience-driven procedural content generation. *IEEE Transactions on Affective Computing*, 2(3):147–161, 2011.