

On Linking Level Segments

Colan F. Biemer
Northeastern University
biemer.c@husky.neu.edu

Seth Cooper
Northeastern University
se.cooper@northeastern.edu

Abstract—An increasingly common area of study in procedural content generation is the creation of level segments: short pieces that can be used to form larger levels. Previous work has used concatenation to form these larger levels. However, even if the segments themselves are completable and well-formed, concatenation can fail to produce levels that are completable and can cause broken in-game structures (e.g. malformed pipes in *Mario*). We show this with three tile-based games: a side-scrolling platformer, a vertical platformer, and a top-down roguelike. To address this, we present a Markov chain and a tree search algorithm that finds a link between two level segments, which uses filters to ensure completable and unbroken in-game structures in the linked segments. We further show that these links work well for multi-segment levels. We find that this method reliably finds links between segments and is customizable to meet a designer’s needs.

I. INTRODUCTION

The field of procedural content generation (PCG) researches algorithms that auto-generate content. A major focus of the field is the creation of levels [1]. Approaches vary from graph grammars [2] to machine learning [3] to reinforcement learning [4] to behavior trees [5] to search [6] and more. Generators generate whole levels or smaller segments that are combined to form larger levels. Segment generation is a promising area because it has potential to give designers more fine-grained control. To better understand this claim, consider the task of building a *Mario* level with many jumps at the start, then a few enemies, then no jumps, and, finally, many jumps and many enemies. If a generator has to build the described level, then it has a large challenge due to the size of the search space. It can be more manageable to break the designer’s requirements into parts by building smaller segments and combining them together. Unfortunately, little work has studied how to combine level segments to build full levels. This is important because the field works with games more complex than *Mario*, and, as we will show, current approaches are insufficient for more complex games, and we need a better method.

The simplest approach to link level segments is *concatenation*, or simply placing one segment after another [7]. Concatenation has no notion of completable or in-game structures (e.g. doors in *Kid Icarus*). Another approach is to require padding on either side of a generated segment [8]; this implicitly handles the problem of broken in-game structures. Green et al. extend this by using an agent to validate that a level is playable before showing it to the user [8]. However, these examples only use *Mario*.

As an alternative to concatenation, we present the process of *linking* for 2D tile-based games, which creates a *linker*—a tiny level segment that connects two level segments. The full level is the concatenation of the first segment, the linker, and the second segment. The full level should be *completable* (i.e. a player can get from the beginning to the end) and *unbroken* (i.e. a level should contain no broken in-game structures). A linker is *usable* if it results in a level that is completable and unbroken.

To generate usable linkers,¹ we present a method of linking which uses two Markov chains and a tree search. This method relies on breaking a set of input levels into vertical or horizontal level slices. These level slices are used as input for two Markov chains. Both chains are used for structure completion for their respective level segment to guarantee an unbroken level by adding level slices to the linker—a limitation is that Markov chains can fail for unseen input. If we concatenate with the linker generated by the Markov chains, we only have a guarantee that a generated linker will result in an unbroken level. To address completable, we use a tree search which runs on the input level slices or a set defined by the designer. The output of the search is an ordered list of level slices—the list may be empty. These are added in between any level slices added by the two Markov chains to form the linker. If a linker can be produced, it is guaranteed to be usable.

We test with three games: *Mario*, a horizontal platformer; *Kid Icarus*, a vertical platformer; and *DungeonGrams*, a top-down roguelike.² For all three games, we test linking with level segments previously generated by Gram-Elites in past work [9].³ We choose *Mario* as a baseline because much research in procedural content generation via machine learning [3] for games uses it. *Icarus* is a platformer but has more complicated vertical pathing than *Mario*. *DungeonGrams* is a different genre, which means it has distinct requirements and can show that our method works for more than just platformers.

We compare concatenation with our linking approach. Linkers always result in unbroken levels. We also find that linking is always able to find usable linkers between two segments for *Mario* and *DungeonGrams*. Concatenation works well for *Mario*, but often fails for *DungeonGrams*—structures are broken, the level is not completable, or both. For *Kid Icarus*, our approach almost always finds usable linkers and concatenation

¹<https://github.com/bi3mer/LinkingLevelSegments>

²<https://github.com/crowdgames/dungeongrams>

³<https://github.com/bi3mer/GramElitesData>

is likely to fail. We extend testing linkers between only two segments by using them to form larger levels consisting of multiple linked segments. For *DungeonGrams*, we find that increasing the number of segments can result in levels that are not usable. For *Mario* and *Icarus*, though, increasing the number of segments does not affect usability.

II. RELATED WORK

This section is broken into four parts. First, we look at previous work that has combined vertical slices and tree search to generate levels. Second, we consider past work which connects dungeon rooms to form larger dungeons. Third, we discuss the similarity of our work to level repair and the work done in the area. Lastly, we examine previous work that has combined level segments to form larger levels.

A. Markov Chains and Tree Search

2D tile-based game levels can be broken down from a grid to an ordered set of horizontal or vertical slices. Dahlskog et al. [10] use these slices as input to an n-gram [11] to generate levels. Summerville et al. expand on this by using slices as input for a Markov chain where the output can be more than one level slice [12]. They use Monte-Carlo tree search (MCTS) [13] to generate full levels. This addresses a weakness of generation by following probabilities in the Markov chain: there is no guarantee on what will be produced. The addition of tree search allows Summerville et al. to produce levels that are completable and unbroken. Further, they guide the search to match target characteristics.

B. Connecting Dungeon Rooms

Liapis [14] uses evolution to generate a higher-order representation of a dungeon based on rooms. A secondary evolutionary process is run to build these rooms where every room has constraints (e.g. must have a connection to the left and right). These constraints are how rooms, or level segments, are correctly linked to form a larger level.

An alternative approach comes from Ashlock and McGuinness [15]. They use evolution to fill in a grid with different tile types. Dynamic programming is built into the fitness function to guarantee that a path exists between all checkpoints in the level, which ensures that the larger level will be completable. They then use a room membership algorithm to find the rooms that were evolved and a separate algorithm to find adjacent rooms. Lastly, they fill in the rooms with content based on required and optional content.

C. Level Repair

One way to view this work is that we are proposing a repair agent for a level formed from level segments. Cooper and Sarkar propose one approach to level repair using a pathfinding agent that can find a path that is impossible following the game’s mechanics [16]. Points in the path that are impossible are repaired by, for example, removing a solid block that the agent jumped through.

Another approach comes from Zhang et al. [17], where they use a GAN [18] to generate a level, but find that it

does not reliably encode playability. To address this they use a “generate-then-repair” framework. Levels are generated with the GAN, and then repaired with mixed integer linear programming, which finds the minimum number of changes that result in a fully playable level.

Jain et al. [19] use an autoencoder trained on *Mario* levels from the VGLC [20]. Among other uses, the autoencoder network is used to repair levels. In their work, a level is broken into windows or small level segments which can be input into the network. By placing a broken window into the network, they show that the network will output a similar segment that is playable.

D. Using Level Segments

Volz et al. [7] trained a generative adversarial network (GAN) [18] on a single *Mario* level. Their innovation in the process was to search the latent space of the generative network with CMA-ES [21]. In one of their experiments, they tested forming larger levels with increasing difficulty. This was accomplished by generating the level segments with a target difficulty and then concatenating them together, which means there is no guarantee that the larger level is unbroken or completable.

Sarkar and Cooper [22] use variational autoencoders (VAE) [23] trained on levels with path information to generate level segments. The VAE was trained to encode a level segment and then use the encoding to predict the segment that follows. A level is generated by starting with one segment and then sequentially building new segments with the VAE. Meaning, the VAE is performing the role of generating segments sequentially in a way that they connect to each other.

Green et al. [8] start with a corpus of *Mario* levels that have been generated with a required padding of two vertical slices of ground tiles on either side [24]. The segments are linked together with the padding to form larger levels. Larger level sequences are found by using FI-2Pop [25]. Infeasible sequences are optimized to be completable. Feasible sequences are optimized to match a target mechanics sequence. The padding guarantees unbroken levels and FI-2Pop finds combinations of segments that are completable.

Li et al. [26] use an ensemble of Markov chains to generate *Mega Man* levels. They use a first-order Markov chain to model the direction (i.e. horizontal or vertical) of a set of rooms. They use two L-shape Markov chains [27] to generate the rooms, one for horizontal rooms and the other for vertical rooms. They place these rooms together and run a check to guarantee that a path from one to the other exists. If a valid path does not exist, they re-sample until two rooms with a connection are found, guaranteeing a completable level.

III. APPROACH

Here, we describe our approach. First, we give a brief overview of Gram-Elites, which previously generated the level segments we use. Second, we describe our approach to link two segments together. Finally, we review the games used in this work.

```

procedure BUILD_LINK(start, end)
  for s in forward_chain.get(start) do
    for e in back_chain.get(end) do
      m ← tree_search(start + s, e + end)
      if m ≠ None then
        return s + m + e
  return None

```

Fig. 1. A pseudocode description for building a link.

A. Gram-Elites

Gram-Elites [9] is an extension to MAP-Elites [28] that uses n -grams for population generation and the genetic operators mutation and crossover. These operators use a concept called *connection*, which runs a breadth-first search through the n -gram to ensure that post-modification the new segment is generable by the n -gram. Like MAP-Elites, Gram-Elites uses behavioral characteristics [29] to differentiate level segments—we give a brief description of the behavioral characteristics used for each game in section III-C. Gram-Elites guarantees that all output segments are generable by an n -gram and optimizes segment completability.

For this work, we filter out segments that were not completable.

B. Generating Linkers

When building a linker between two segments, there are three problems. First, does the starting segment end with an incomplete structure? (Figure 2a is an example of this problem.) Second, does the end segment begin with an incomplete structure? Third, does the linker result in a completable level when used to link the two segments? We start by addressing the first two problems.

One option to address level segments that start or end with incomplete structures is to filter them out, but we believe a more flexible approach is called for. We address this with structure completion. In some games, like *Mario*, it would be simple to check for incomplete pipes at the edge of segments and add a custom level slice to complete them. To be more general, we take advantage of using levels that are generable by an n -gram. We use two Markov chains to address incomplete structures. (Note that these chains do not modify the level segments themselves and do not act as level repairers.) The first, the forward chain, addresses potentially incomplete structures at the end of the starting segment. This chain is initialized from data in the direction the designer meant (e.g. *Mario* is from left to right). The second, the back chain, reads in reverse. We filter the data each Markov chain receives to only include level slices with tiles associated with game structures, resulting in input and output that can be greater than one level slice. Further, input is limited by the size of a structure—e.g. *DungeonGrams* structures are four columns long—to prevent the potential problem of adding additional structures where only one is called for.

Figure 2 shows the forward chain in action. It uses the last slice of the starting segment as input into the chain. If there

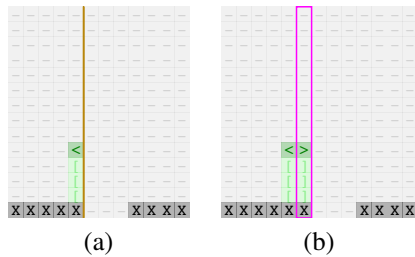


Fig. 2. (a) The tan line marks where the starting segment ends and the end segment begins. There is an incomplete pipe with concatenation. (b) The magenta box shows the output of the forward chain which completes the pipe.

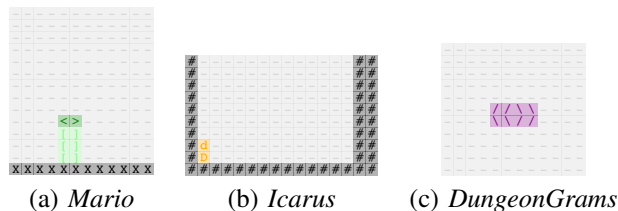


Fig. 3. Examples of structures for all three test games. *Mario* has pipes which are two columns wide and some number of rows tall. *Icarus* has doors that are one column wide and two rows tall. *DungeonGrams* has a structure that is 4 columns wide and 2 rows tall.

is no output, there are no incomplete structures at the end of the starting segment. If there is output, check if there is output for the last two slices. If so, check for three, and so on until there is no output. We use the output associated with the most slices in the Markov chain. The back chain follows the same process, but in the opposite direction. Once both chains run, we have a guarantee that the concatenation of the starting segment, forward link, back link, and end segment has no incomplete structures if the input segments are generable by the n -gram used in Gram-Elites, which comes from the data used to create the forward and backward Markov chains. If either chain receives a structure not present in the input data, structure completion with Markov chains will fail.

The next step is to address the problem of guaranteeing that a linked level is completable. *Linking slices* are level slices which do not contain in game structures. By default, these are the level slices not input into the previously described Markov chains. Alternatively, the designer can define them. A breadth-first search uses linking slices to find the minimum number of slices required to make the fully linked level completable. The output can be zero slices, but we can add constraints (e.g. require at least k slices). We use a max-depth check to prevent an infinite search. A benefit of using designer-defined linking slices is that the search space is reduced and the algorithm is much faster.

View figure 1 for pseudocode of the proposed method. The for loops for both chains allow for the possibility that there may be multiple ways to complete a structure.

C. Games

Here we review the three games used to evaluate linking. When evaluating completability, we pad the beginning and end of levels being evaluated with a few level slices. This allows

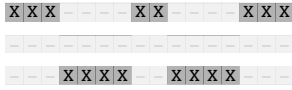


Fig. 4. The three linking slices used for for *Kid Icarus*.

us to easily define where the agent starts and where the agent must reach for the level to be classified as completable.

Mario is a horizontal platformer. As input for the linker structures, we use training levels from the VGLC [20]. The level segments previously built with Gram-Elites use the same input levels, and the behavioral characteristics are linearity and leniency, which are the axes of the MAP-Elites grid. For the linking completable tree search, default linking slices were used with no designer intervention. As a reminder, the set of default linking slices does not include level slices with tiles related to pipes. The max depth for the linking tree search is 7. For our agent, we used a modified Summerville A* agent [20] to calculate the furthest point in the level that can be reached. The only in-game structures are pipes, see Figure 3a.

Kid Icarus is a vertical platformer. As input for the linker structures, we use training levels from the VGLC [20]. The level segments previously built with Gram-Elites use the same levels. The axes of the grid are density and leniency. For the linking completable tree search, we use one empty row and two rows with platforms designed to connect to anything else, see figure 4. The max depth for the linking tree search is 7. For our agent, we use a modified Summerville A* agent [20] to calculate the furthest point in the level that can be reached. *Kid Icarus* has one structure which is a door, see figure 3b.

DungeonGrams is a top-down roguelike developed to evaluate Gram-Elites. The player has to progress from left-to-right and hit switches to unlock a portal and then use it as an exit to beat a level. There are spikes and enemies that the player has to avoid while navigating a level. *DungeonGrams* also has as stamina mechanic where the player has a limited number of moves before they lose. The player can increase their stamina by finding food in the level.

As input for the linker structures, we use the 44 training levels made for Gram-Elites. The axes of the grid are density and leniency. For the linking completable tree search, we use three columns with food: bottom, middle, and top of each respective column. To evaluate completable, we use a search-based agent already built for *DungeonGrams*. The max depth of the linking tree search is 4, which is lower than the other two games due to the agent being slower as a result of a larger search space. *DungeonGrams* also has one structure that can be seen in figure 3c.

IV. EVALUATION

We break our evaluation into two parts. First we evaluate the effectiveness of concatenation and compare it to our linking

	Type	Unbroken	Completable	Usable
<i>Mario</i>	Concatenation	0.97	0.92	0.89
	Linking	1.00	1.00	1.00
<i>Icarus</i>	Concatenation	0.96	0.11	0.11
	Linking	1.00	0.99	0.99
<i>DG</i>	Concatenation	0.52	0.58	0.23
	Linking	1.00	1.00	1.00

TABLE I
PERCENTAGE OF LINKS THAT RESULTED IN UNBROKEN LEVELS, COMPLETEABLE LEVELS, AND USABLE LEVELS. DG IS SHORT FOR *DungeonGrams*.

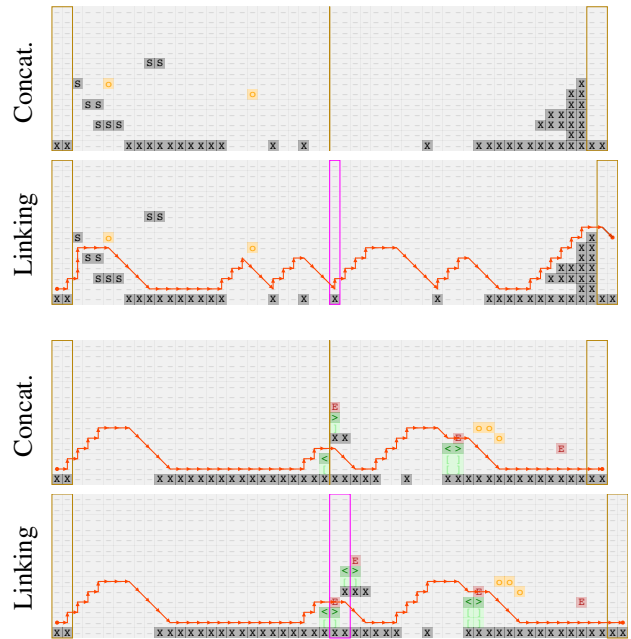


Fig. 5. Two examples for *Mario*. The example on top shows concatenation failing due to a large gap. The bottom displays the longest linker produced. The tan line in the middle is where the link would have been placed for concatenation. The tan boxes on the left and right show the padding. The magenta box in the middle shows the linker found. The red lines show a path through the level if it was beatable.

algorithm. Second, we examine how well linkers made for two segments extend to linking multiple segments.

A. Concatenation Versus Linking

Segments generated by Gram-Elites are organized by the MAP-Elites grid where each axis represents a behavioral characteristic [29]. Rather than selecting random segments and linking them together, we use the grid to select similar segments by using neighboring segments. There are 13,182 possible links for *Mario*, 9,453 for *Kid Icarus*, and 6,086 for *DungeonGrams*.

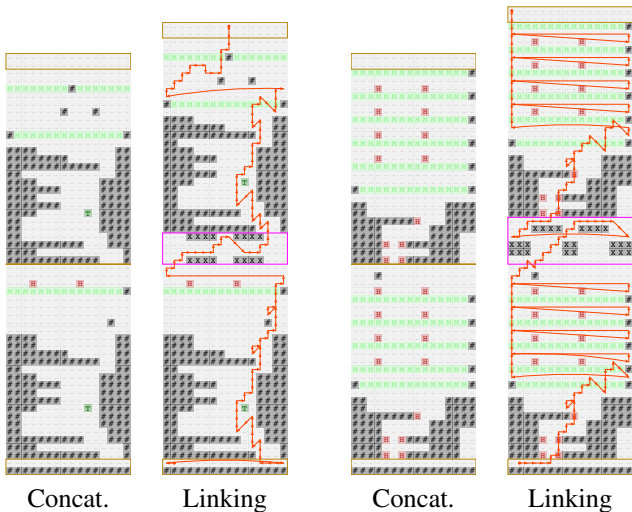


Fig. 6. Two examples for *Icarus*. Both show concatenation failing due there being no possible path to complete the level. The right example of concatenation and linking shows the longest linker required to make a level completable. Padding is shown with the tan boxes at the top and bottom of the level. The tan line in the middle is where the link would have been placed for concatenation. The magenta box shows the linker found. The red lines show a path through the level if it was beatable.

Table I shows the percentage of concatenations and links found that are unbroken, completable, and usable (i.e. both unbroken and completable). With *Mario*, we can see that concatenation is 0.92 likely to result in a level that is completable by an agent and 0.89 likely to result in a level that is usable, showing that concatenation is likely to be an effective strategy for most use cases in a simple platformer. Linking is successful for every possible link. Figure 5 shows two examples where concatenation fails. The top example shows that concatenation can fail when a gap is too large. The bottom example shows the largest linker required for *Mario*. The linker completed two structures, but did not affect completable as the concatenated version was also completable.

Concatenation performs well for *Icarus* in terms of producing an unbroken level. However, it performs the worst at generating a completable level at just eleven percent. Linking, though, finds a usable link for almost every single starting and ending segment given. Figure 6 shows two examples where concatenation cannot produce a completable level. In both cases the agent does not have enough space to make the required jump. Tree search resolves this by adding two platforms for the agent. The right-most example shows the largest linker built.

Concatenation performs poorly for *DungeonGrams*, and is unlikely to produce a usable level. This is due to two reasons. First, the size and complexity of the structures makes it unlikely for two segments to line up if either or both level segments have unfinished structures. Second, if the two segments do not have enough food, the concatenated level is impossible to beat. Linking, though, handles both problems and always produces a usable link. Figure 7 shows two examples where concatenation fails. In both cases, structures are incomplete

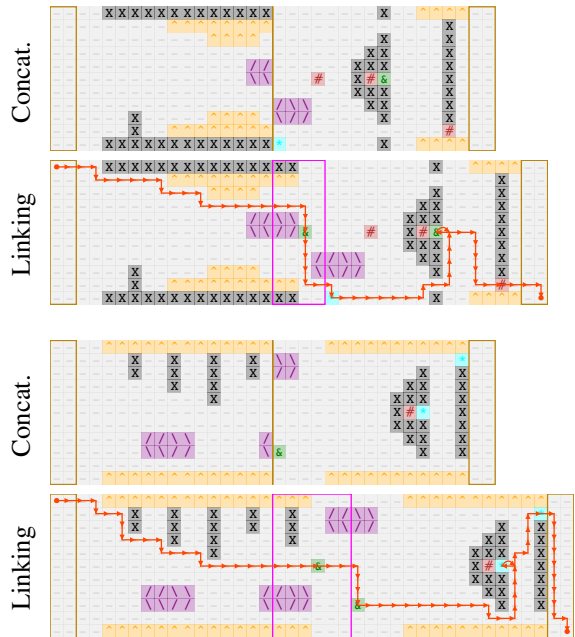


Fig. 7. Two examples for *DungeonGrams* where concatenation fails due to the agent running out of stamina. (& represents food.) The bottom concatenation and linking example shows the largest linker found. The tan boxes on the left and right show the padding. The tan line in the middle is where the link would have been placed for concatenation. The magenta box shows the linker found. The red lines show a path through the level if it was beatable

Game	Mean Length	Median Length	Max Length	Mean D_{BC}	Median D_{BC}	Max D_{BC}
<i>Mario</i>	0.110	0	2	0.001	0.000	0.020
<i>Icarus</i>	2.658	3	6	0.008	0.008	0.029
<i>DG</i>	1.147	1	6	0.012	0.011	0.061

TABLE II

FOR EACH GAME, SHOWS THE MEAN, MEDIAN, AND MAX FOR LENGTHS OF LINKERS FOUND AND D_{BC} . DG IS SHORT FOR *DungeonGrams*.

and there isn't enough food for the agent. Linking succeeds by completing the structures with the forward and back chains, and tree search adds food in the middle to make the level completable. The bottom example shows the largest linker found for *DungeonGrams*.

The usability of a linker isn't the only important metric to consider. Ideally, a linker should not be identifiable or distracting to the player (e.g. the same set of level slices between every segment). To examine this, we report on the lengths of linkers found and the change in behavioral characteristics of a two segment level when a linker is used. We determine the latter by finding the behavioral characteristics (BCs) of the concatenated level and the linked level, and calculate the euclidean distance between the two. The result is D_{BC} .

The results can be seen in Table II. For *Mario*, we can see that a linker tends not to be required and, when it is, the worst case is a linker with two level slices. As a result, D_{BC} must be small. To put the value in context, the implementation of Gram-Elites tessellates the grid by changes in BC every 0.025 for *Mario*. For *Mario*, if a linked level and its corresponding

Segments	<i>Mario</i>	<i>Icarus</i>	<i>DG</i>	<i>DG-Food</i>
2	1.00	1.00	1.00	1.00
3	1.00	1.00	0.99	0.99
4	1.00	1.00	0.98	0.99
5	1.00	1.00	0.97	0.99

TABLE III

SHOWS THE PERCENTAGE THAT LINKED LEVELS ARE USABLE. *DG* IS SHORT FOR *DungeonGrams* WHERE AN EMPTY LINKER IS ALLOWED. *DG-Food* REQUIRES THAT THE TREE SEARCH ADDS AT LEAST ONE LEVEL SLICE TO THE LINKER.

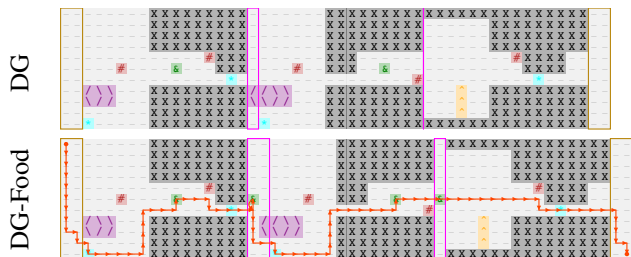


Fig. 8. Example of three segments being linked in *DungeonGrams*. Tan boxes on either side represent padding. Magenta boxes in the middle represents a linker and the magenta line shows that no linker was necessary for the second and third level segment for linking (*DG*). The agent fails to reach the end for *DG* due to a lack of food, but this was not a problem for linking with food required (*DG-Food*).

concatenated level have a D_{BC} smaller than 0.025, then we consider them to be very similar. *Icarus* is different in that it almost always requires a linker with two to three linking slices. The max length is also much longer than *Mario*. The increment for *Icarus*'s grid is 0.0125. In this case, the mean and median of D_{BC} show that most linkers are unlikely to result in large changes to the BCs. The max D_{BC} shows that the worst case is a jump of two to three neighbors in the grid. Finally, *DungeonGrams* on average has smaller linkers than *Icarus* but has the same max linker length. The mean and median of D_{BC} is quite small: the Gram-elites grid is 0.05 per bin in *DungeonGrams*.

Overall, linking usually results in minimal modifications to the player's experience per segment while providing the benefit of completability and unbrokenness.

B. Linking Multiple Segments

In this section, we test the linkers made for two segments by combining multiple segments with them. To test, for various numbers of segments, we generated one thousand random levels by following the MAP-Elites grid, neighbor by neighbor. Linkers were selected only if they had a playability guarantee for the starting and ending level segment. The results are in Table III. Notably, we see that all two-segment levels are usable, as should be the case using the linkers.

For both platformers, we can see that increasing the number of segments does not change the results: all the linked multi-segment levels were usable.

For *DungeonGrams*, we updated linking for a new approach. *DG* refers to when an empty link is allowed. *DG-Food* forces

the link to have at least one linking slice from the tree search, meaning there is at least one column with food in the link since we use custom level slices which always have food. We find that *DG* performs increasingly poorly when compared to the platformers as the number of segments increases. In contrast, the simple change for *DG-Food* results in larger levels almost always being fully playable. Figure 8 shows an example where *DG* fails and *DG-Food* succeeds. Failure occurs since linking guarantees that a path exists between two segments, but does not ensure that the player will have the same stamina at the start of a new segment as they do when starting the game. While food in *DungeonGrams* does not give the player back their max stamina, it does increase the likelihood that the agent can make it through the next segment.

V. CONCLUSION

In this work, we present an alternative to concatenation for linking level segments of 2D tile-based games. Our approach uses two Markov chains and a tree search to guarantee that any linker produced results in a level that has no unbroken structures and is completable by an agent. We test on three games: *Mario*, *Kid Icarus*, and *DungeonGrams*.

When linking two segments, we find that our approach outperforms concatenation for all three games. In the two-segment case, it always finds a usable linker for both *Mario* and *DungeonGrams*. For *Kid Icarus*, linking works for the vast majority of segments but can fail. On examining the failure cases, we believe this is partly due to the agent not perfectly playing the game by not correctly modelling blocks where the player can jump through and land on the top. We intend to improve the agent as part of future work.

We extend our evaluation and test how pre-built two-segment linkers work with sequences of level segments greater than two. We find that for *Mario* and *Kid Icarus*, all multi-segment levels are usable. For *DungeonGrams*, though, performance is worse due to the stamina mechanic. We address this by requiring the link to always contain at least one column with food, and find that almost every large level can be completed. Another approach is to extend the linking strategy to link all k segments simultaneously, where we currently only solve for $k = 2$.

As it stands, our approach to linking can work with any segment-based generator if it generates segments that are generable by an n -gram; in this work, linkers used the same datasets that were used to generate the segments. As future work, we believe it is important to make a linking method that works for all generators, which means addressing the limitation of requiring n -gram generability. At the minimum, this requires updating structure generation to correctly handle unseen in-game structures. To further extend this work, we want to use the directed graph we built to evaluate linking, and use it to build larger levels tailored to individual players.

REFERENCES

- [1] N. Shaker, J. Togelius, and M. J. Nelson, *Procedural content generation in games*. Springer, 2016.

- [2] E. Hauck and C. Aranha, "Automatic generation of Super Mario levels via graph grammars," in *2020 IEEE Conference on Games (CoG)*. IEEE, 2020, pp. 297–304.
- [3] A. Summerville, S. Snodgrass, M. Guzdial, C. Holmgård, A. K. Hoover, A. Isaksen, A. Nealen, and J. Togelius, "Procedural content generation via machine learning (pcgml)," *IEEE Transactions on Games*, vol. 10, no. 3, pp. 257–270, 2018.
- [4] A. Khalifa, P. Bontrager, S. Earle, and J. Togelius, "PCGRL: Procedural content generation via reinforcement learning," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 16, no. 1, 2020, pp. 95–101.
- [5] A. Sarkar and S. Cooper, "Procedural content generation using behavior trees (PCGBT)," *arXiv preprint arXiv:2107.06638*, 2021.
- [6] J. Togelius, G. N. Yannakakis, K. O. Stanley, and C. Browne, "Search-based procedural content generation: A taxonomy and survey," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 3, no. 3, pp. 172–186, 2011.
- [7] V. Volz, J. Schrum, J. Liu, S. M. Lucas, A. Smith, and S. Risi, "Evolving Mario levels in the latent space of a deep convolutional generative adversarial network," in *Proceedings of the genetic and evolutionary computation conference*, 2018, pp. 221–228.
- [8] M. C. Green, L. Mugrai, A. Khalifa, and J. Togelius, "Mario level generation from mechanics using scene stitching," in *2020 IEEE Conference on Games (CoG)*. IEEE, 2020, pp. 49–56.
- [9] C. Biemer, A. Hervella, and S. Cooper, "Gram-elites: N-gram based quality-diversity search," in *Proceedings of the FDG workshop on Procedural Content Generation*, 2021, pp. 1–6.
- [10] S. Dahlskog, J. Togelius, and M. J. Nelson, "Linear levels through n-grams," in *Proceedings of the 18th International Academic MindTrek Conference: Media Business, Management, Content & Services*, 2014, pp. 200–206.
- [11] D. Jurafsky, *Speech & language processing*. Pearson Education India, 2000.
- [12] A. Summerville, S. Philip, and M. Mateas, "MCMCTS PCG 4 SMB: Monte carlo tree search to guide platformer level generation," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 11, no. 1, 2015.
- [13] R. Coulom, "Efficient selectivity and backup operators in monte-carlo tree search," in *International conference on computers and games*. Springer, 2006, pp. 72–83.
- [14] A. Liapis, "Multi-segment evolution of dungeon game levels," in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2017, pp. 203–210.
- [15] D. Ashlock and C. McGuinness, "Automatic generation of fantasy role-playing modules," in *2014 IEEE Conference on Computational Intelligence and Games*, 2014, pp. 1–8.
- [16] S. Cooper and A. Sarkar, "Pathfinding agents for platformer level repair," in *AIIDE Workshops*, 2020.
- [17] H. Zhang, M. Fontaine, A. Hoover, J. Togelius, B. Dilkina, and S. Nikolaidis, "Video game level repair via mixed integer linear programming," *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 16, no. 1, pp. 151–158, Oct. 2020, number: 1. [Online]. Available: <https://ojs.aaai.org/index.php/AIIDE/article/view/7424>
- [18] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [19] R. Jain, A. Isaksen, C. Holmgård, and J. Togelius, "Autoencoders for level generation, repair, and recognition," in *Proceedings of the ICCG workshop on computational creativity and games*, vol. 9, 2016.
- [20] A. J. Summerville, S. Snodgrass, M. Mateas, and S. Ontanón, "The VGLC: The video game level corpus," *arXiv preprint arXiv:1606.07487*, 2016.
- [21] N. Hansen, S. D. Müller, and P. Koumoutsakos, "Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES)," *Evolutionary computation*, vol. 11, no. 1, pp. 1–18, 2003.
- [22] A. Sarkar and S. Cooper, "Sequential segment-based level generation and blending using variational autoencoders," in *International Conference on the Foundations of Digital Games*, 2020, pp. 1–9.
- [23] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," *arXiv preprint arXiv:1312.6114*, 2013.
- [24] A. Khalifa, M. C. Green, G. Barros, and J. Togelius, "Intentional computational level design," in *Proceedings of The Genetic and Evolutionary Computation Conference*, 2019, pp. 796–803.
- [25] S. O. Kimbrough, G. J. Koehler, M. Lu, and D. H. Wood, "Introducing a feasible-infeasible two-population (FI-2Pop) genetic algorithm for constrained optimization: Distance tracing and no free lunch," *European Journal of Operational Research*, 2005.
- [26] B. Li, R. Chen, Y. Xue, R. Wang, W. Li, and M. Guzdial, "Ensemble learning for Mega Man level generation," *arXiv preprint arXiv:2107.12524*, 2021.
- [27] S. Snodgrass and S. Ontanón, "A hierarchical approach to generating maps using markov chains," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 10, no. 1, 2014.
- [28] J.-B. Mouret and J. Clune, "Illuminating search spaces by mapping elites," *arXiv preprint arXiv:1504.04909*, 2015.
- [29] G. Smith and J. Whitehead, "Analyzing the expressive range of a level generator," in *Proceedings of the 2010 Workshop on Procedural Content Generation in Games*, 2010, pp. 1–7.