

Medida de similitud para experiencias de juego basada en el ritmo

José A. Torres-León¹, Marco A. Moreno-Armendáriz¹,
Carlos A. Duchanoy², Hiram Calvo¹

¹ Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Laboratorio de Ciencias Cognitivas Computacionales,
México

² Gus Chat,
México

{jtorres12019,mam_armendariz,hcalvo}@cic.ipn.mx,
carlos.duchanoy@gus.chat

Resumen. En este escrito se presenta una medida de similitud para evaluar la experiencia de juego que brinda una sección de nivel. La medida es una propuesta para evaluar matemáticamente el concepto de “ritmo” en los video juegos. Se presenta una investigación sobre medidas existentes para evaluar las experiencias de juego en el área de la generación procedimental de contenido y por qué la medida de similitud presentada es relevante. Además, se presenta un estudio realizado sobre la utilidad de la medida para al menos tres diferentes experiencias de juego.

Palabras clave: Inteligencia artificial, medidas de similitud, experiencia de juego, generación procedimental de contenido.

Rhythm Based Similarity Measure for Game Experiences

Abstract. In this paper a similarity measure for evaluating game experiences in a level chunk is presented. The measure is a proposal to mathematically evaluate the “rhythm” concept on videogames. Research on existing measures to evaluate the game experiences in the procedural content generation area is presented and why the proposed similarity measure is relevant. In addition, a utility study of the measure to at least three different game experiences are presented.

Keywords: Artificial intelligence, similarity measures, game experience, procedural content generation.

1. Introducción

La generación procedimental de contenido (GPC) es una disciplina que se concentra en la generación algorítmica de contenido, principalmente, para videojuegos. La GPC puede usarse para generar todo tipo de contenido, se ha usado para crear niveles, música, sonidos, flujos de misiones, personajes e incluso reglas de juego, ya sea en tiempo de ejecución o como herramienta de diseño [1]. El alto impacto que ha tenido esta disciplina ha generado un gran interés en la comunidad académica y ha desembocado en un considerable número de libros y estudios en las últimas dos décadas [2, 3, 4, 5, 6, 7, 8].

En general, estos métodos requieren de una evaluación cuantitativa de la calidad o las características del contenido generado, de manera que se pueda afirmar si los resultados que de ellos se obtienen son adecuados. Por ello, existen una gran variedad de formas de evaluar el contenido de los videojuegos.

La propuesta que se presenta, consiste en una nueva medida de similitud para experiencias de juego, la cual toma como base la idea del *ritmo*, que ha sido usado por los diseñadores de niveles de video juegos en la industria. Esta medida de similitud pretende ser una herramienta de apoyo para sistemas de GPC y de inteligencia artificial (IA) involucrados en las tareas de desarrollo de videojuegos.

La propuesta, surge al detectar la necesidad de los sistemas de generación automática de niveles para determinar si el contenido generado incluye una cierta experiencia de juego.

Por ello, en la siguiente sección se presentan algunos trabajos del estado del arte sobre la evaluación del contenido de los videojuegos, revisando qué aspectos se consideran en general al evaluar el contenido y cuáles son las propuestas sobre la evaluación de experiencias de juego.

2. Métodos de evaluación de contenido de videojuegos

Los primeros acercamientos a la evaluación de contenido de videojuegos se basan en la capacidad de un experto para asegurar la calidad del producto. Por ello, los primeros esfuerzos para formalizar la evaluación de contenido se centran en traducir y/o resumir el conocimiento de un experto en conjuntos de consejos y buenas prácticas para que los no expertos sean capaces de evaluar el contenido de un videojuego de manera similar a como lo haría un experto.

Algunos ejemplos de las propuestas más tradicionales se encuentran en [10, 11], que proporcionan una serie de heurísticas o consejos, que servirán para determinar la calidad del nivel y la experiencia de juego que brindan. Estas propuestas ofrecen una primera noción acerca de lo que debería considerar cuando se evalúan contenido para un videojuego. Las siguientes son dos de las recomendaciones más relevantes de [10]:

1. “El *gameplay* (la *jugabilidad*) debe ser balanceado, con múltiples caminos para ganar”. Para medir este aspecto se puede contar el número de caminos diferentes en los niveles para llegar a la plataforma meta,
2. “La fatiga del jugador se minimiza al variar las actividades y el paso (ritmo) durante el juego”. Para medir la variación en las actividades se puede implementar una medida de ritmo.

Por su parte, en [11] se afirma que cada género de videojuego tiene un ritmo característico de la dificultad, este ritmo o paso se ve reflejado al momento de jugar. Por ejemplo, en un juego de plataforma es común que entre un reto y otro el nivel ofrezca un espacio para descansar al jugador o que se ofrezcan dos rutas distintas, una llena de enemigos y recompensas y otra que es un camino vacío y que no presenta ni enemigos ni recompensas. Estos cambios a lo largo del nivel marcan una forma específica de interactuar con él, misma que implica subidas y bajadas en la dificultad que implica superar cada sección del nivel para el jugador, estos cambios son los que constituyen el ritmo característico. En conjunto, estas referencias dejan en claro que características como la variedad de caminos, la dificultad o el tipo de reto presente en ellos, así como el ritmo de estas características en los caminos, son útiles para definir el tipo de nivel al que se enfrenta un jugador.

2.1. Modelado de la experiencia de juego

Una de las motivaciones principales para el desarrollo de las métricas para niveles de videojuegos es que describan las diferentes experiencias de juego.

Como se indica en [12], la experiencia de juego se refiere a la relación que existe entre las reacciones del jugador al momento de interactuar con el videojuego y el contenido con el que interactúa en dicho momento, por ello se considera como una función del contenido del juego y el jugador. Por su parte, la *jugabilidad* del videojuego se refiere a cómo se ha diseñado el contenido para que el jugador interactúe con él.

Entonces, el contenido se diseña de modo que hay una cierta *jugabilidad* asociada a éste. Sin embargo, la experiencia de juego varía dependiendo de cada jugador dada una misma *jugabilidad*, por lo tanto, aunque no se puede asegurar que todos los jugadores tendrán una misma experiencia bajo una misma *jugabilidad*, sí existe una relación entre la *jugabilidad* y la experiencia de juego, al hacer que todos los jugadores interactúen de una misma forma con el videojuego, se espera que sus reacciones sean similares.

Para evaluar la experiencia de juego existen algunas propuestas que consideran tomar mediciones directamente del jugador durante sesiones de juego y después de ellas, tales como [13, 14, 15, 16, 17, 18, 19, 20], sin embargo, estas evaluaciones son costosas, pues algunas requieren de equipo especializado para tomar las medidas y en general, todas requieren de jugadores que deben ser remunerados por su participación en el proyecto y que inevitablemente dan una evaluación sesgada debida a sus propias habilidades y preferencias de juego.

Sin embargo, también existen trabajos como [21, 22] cuyos autores hacen un análisis del contenido de juego presente en una porción de nivel y lo relacionan con la mecánica o reto que presenta dicha porción de nivel para el jugador.

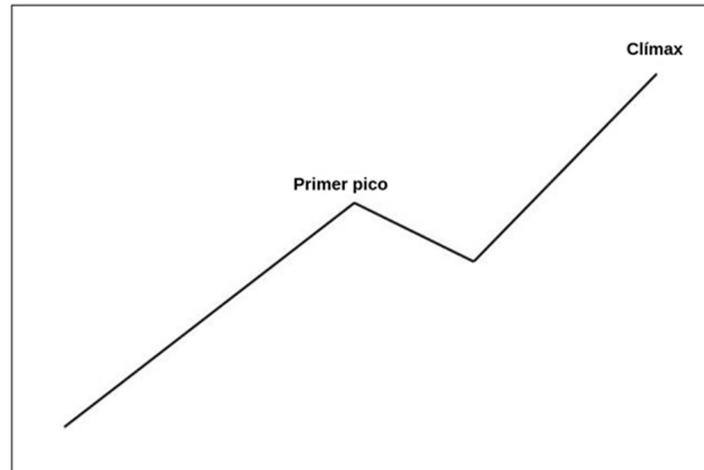


Fig. 1. Ilustración de una característica de nivel de videojuegos con dos fluctuaciones.

Aunque estos trabajos dejan de lado aspectos importantes como la dificultad o el ritmo que evalúan los expertos para determinar si una experiencia de juego está presente en cierto contenido de un videojuego.

3. Medida de similitud de experiencias de juego propuesta

Con base en los conceptos presentados en la sección 2, se ideó una nueva evaluación para experiencias de juego, que toma en consideración el ritmo de las características de un nivel, en otras palabras, esta medida toma en consideración el cambio de una medida a lo largo de un nivel del videojuego, permitiendo analizar el nivel en términos de cómo éste fue diseñado y qué jugabilidad ofrece a lo largo de la ruta. Por ello, se hace la suposición de que en cada experiencia de juego debe existir un ritmo específico en las características de las rebanadas del nivel, lo cual implica que a partir del ritmo de esas características es posible afirmar que una rebanada de nivel ofrece una experiencia de juego.

El ritmo característico de una característica de nivel se ilustra en la Fig. 1, en la cual se observan dos picos o fluctuaciones en la medida de la característica.

La medida de similitud propuesta evalúa el comportamiento de las características medidas del nivel, a partir del uso de métricas que midan características de interés (como la dificultad, la suma de puntos que puede obtener el jugador, la distancia entre dos objetos del nivel, etc.), evaluando estas mediciones a lo largo de un camino que lleva de inicio a fin del nivel. Lo que se busca con esta evaluación es asegurar que el ritmo y el rango de valores (valor máximo y mínimo), de las características a evaluar, son cercanos a un parámetro conocido de una experiencia de juego.

Es decir, nuestra evaluación del cumplimiento de la experiencia de juego se basa en la comparación del nivel medido/evaluado con un nivel ideal/prototipo, donde la

comparación se hace a través del ritmo y el rango de valores de las características de interés.

Por lo tanto, nuestra medida de similitud se divide en dos partes, la “exactitud de ritmo” y la “exactitud de rango de valores”, se definen como “exactitud” debido a que entre más cercanos sean los valores medidos con los valores objetivo, mayor será la certeza de que el nivel evaluado brindará la experiencia de juego deseada. Para fines prácticos, decidimos establecer los valores de similitud en el rango de [0,1], donde ‘0’ es lo más distinto posible y ‘1’ es lo más similar posible.

3.1. Exactitud de ritmo

Para esta parte de la medida de similitud, se utilizó una función con forma de campana, cuyo valor se calcula con base en la Ecuación (1). Para esta función, ‘ x_i ’ es el ritmo de la característica i -ésima, ‘ σ ’ es la varianza deseada del ritmo, que sirve para ampliar o reducir el ancho de la campana, haciendo la evaluación del ritmo más relajada (más amplia) o más estricta (más reducida), ‘ μ ’ es el ritmo objetivo, en la ecuación es la media, indicando que la campana tendrá su centro y, por lo tanto, valor máximo, en este punto:

$$f(x_i) = \frac{1}{\mu\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}. \quad (1)$$

Esta función es normalizada para ajustar sus valores al rango [0,1]. Por lo tanto, la primera función de exactitud para la medida de similitud se muestra en la Ecuación (2):

$$a_r = \frac{f(x_i)}{f(\mu)}. \quad (2)$$

3.2. Exactitud de rango de valores

Para esta parte de la medida de similitud, la meta es evaluar qué tan cercano es el rango de valores medidos contra el rango de valores objetivo. Para este propósito, se proponen a su vez dos diferentes comparaciones entre los rangos. Para definir qué tan iguales son dos rangos, se propone comparar la amplitud del rango y su centro, entre más cercanos sean dos rangos en cuanto a sus centros y sus amplitudes, más semejantes son.

Estas ideas se ilustran en la Fig. 2, donde se muestra una curva de una característica de un nivel de videojuego, cuyos valores son la medida de dicha característica a lo largo de los puntos de un camino de inicio a fin del nivel. En esta Figura se representa el rango objetivo (en azul) en el que esa característica debería estar de acuerdo a la experiencia de juego objetivo y el rango real medido (en verde) de esa característica.

Estos rangos se definen a través de los valores máximo y mínimo, marcados con letras azules.

Sean γ el valor máximo medido, δ el valor mínimo medido, λ el valor máximo objetivo, τ el valor mínimo objetivo, entonces, la primera de estas comparaciones se

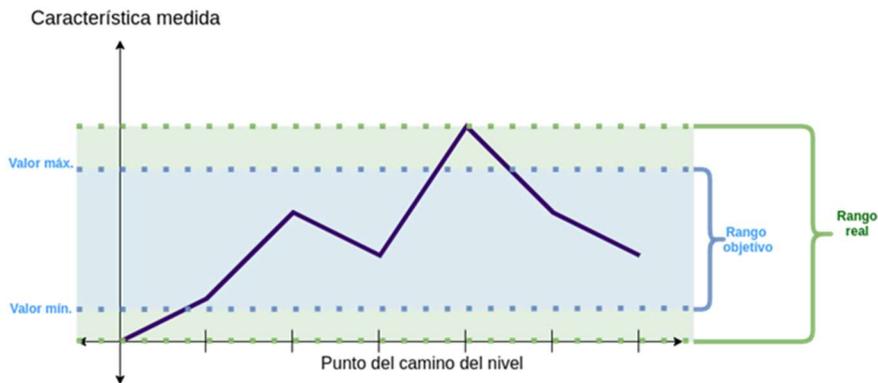


Fig. 2: Ilustración de la exactitud de rango.

muestra en la Ecuación (3), la similitud de amplitud (S_a), que está dada por 1 menos la distancia normalizada entre las amplitudes:

$$S_a = 1 - \frac{|(\gamma - \delta) - (\lambda - \tau)|}{\max(\gamma, \lambda)}. \quad (3)$$

La segunda comparación se muestra en la Ecuación (4), la similitud de centros (S_c), que está dada por 1 menos la distancia normalizada entre los centros:

$$S_c = 1 - \frac{\left(\frac{1}{2}(\gamma - \delta) + \delta\right) - \left(\frac{1}{2}(\lambda - \tau) + \tau\right)}{\max(\gamma, \lambda)}. \quad (4)$$

Estas ecuaciones son válidas si y sólo si $\max(\gamma, \lambda) \neq 0$, de lo contrario, deben utilizarse las Ecuaciones (5) y (6):

$$S_a = 1 - |(\gamma - \delta) - (\lambda - \tau)|, \quad (5)$$

$$S_c = 1 - \left[\left(\frac{1}{2}(\gamma - \delta) + \delta\right) - \left(\frac{1}{2}(\lambda - \tau) + \tau\right) \right]. \quad (6)$$

Finalmente, la exactitud de rango de valores se muestra en la Ecuación (7):

$$a_v = \frac{1}{2}S_a - \frac{1}{2}S_c. \quad (7)$$

Entonces, la medida de similitud S para una característica es una suma ponderada de las exactitudes a_r y a_v , tal y como se muestra en la Ecuación (8), donde $\alpha + \beta = 1$. Los pesos α y β definen qué exactitud es más relevante, la del ritmo o la del rango de valores:

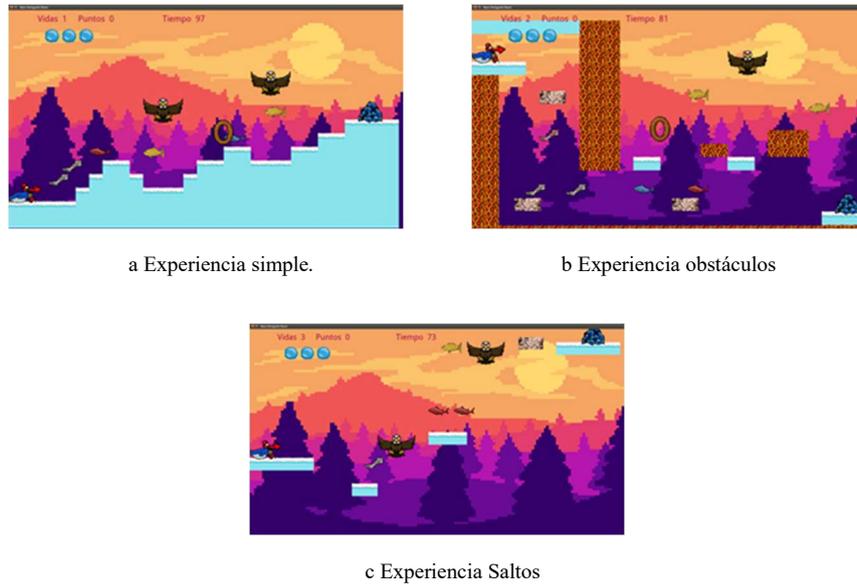


Fig. 3. Experiencias de juego modeladas en Pingu Run.

$$S = \alpha \cdot a_r + \beta \cdot a_v. \tag{8}$$

Dada una experiencia de juego, debe calcularse la similitud de acuerdo a la Ecuación (8) para cada característica de interés y finalmente hacer una suma ponderada de cada una de ellas. La similitud hacia una experiencia de juego se define en la Ecuación (9), donde w_i es el peso de la característica i -ésima y S_i es la función de similitud de esa misma característica, donde $\sum_i w_i = 1$:

$$R_{GX} = \sum_i w_i \cdot R_i. \tag{9}$$

4. Experimentos y resultados

4.1. Niveles prototipo

Para poner a prueba nuestra medida de similitud en un caso de estudio concreto se modelaron tres diferentes experiencias de juego, que se presentan en la Figura 3, para el videojuego “Pingu Run” disponible en [23]. Para tomar las medidas de las características de los niveles usamos el método de evaluación de niveles de videojuegos propuesto en [24], donde se genera un grafo del nivel. En este grafo, cada nodo es un elemento del nivel (piso, enemigos, bonificaciones, etc.) y cada arista entre nodos es una posible transición entre esos bloques.

Este grafo se anota con ocho métricas diferentes: riesgo, obstrucción, distancia, precisión, recompensa de nivel, recompensa de bonus, motivación de nivel y motivación de bonus.

Para obtener una medida del ritmo y de los valores máximo y mínimo, se calculó la ruta de "recompensa de nivel máxima" (la ruta con suma máxima en la medida de recompensa de nivel) con ayuda del algoritmo de Dijkstra para cálculo de rutas óptimas en grafos. Bajo este método de evaluación, el ritmo de una característica se mide contando el número de fluctuaciones que tiene su curva a lo largo de la ruta evaluada. Los valores máximo y mínimo son la medida máxima y la medida mínima de esa misma característica, respectivamente.

Para poder usar la medida de similitud propuesta, es necesario definir los valores objetivo para las ecuaciones presentadas ($\mu, \sigma, \lambda, \tau, \alpha, \beta, w$), puesto que a partir de ellos es que se calcula la similitud contra una experiencia de juego. Como resultado de una definición correcta de estos valores, al evaluar la presencia de una experiencia de juego en una rebanada de nivel, entonces para las rebanadas que sí contengan esa experiencia de juego se debe obtener un valor cercano a 1 y un valor cercano a cero en caso de que esa experiencia de juego no esté presente en la rebanada evaluada.

Para definir estos valores objetivo, se crearon tres niveles para cada una de las experiencias de juego (simple, obstáculos y saltos). Se aplicó la metodología de evaluación seleccionada, obteniendo un grafo por cada nivel evaluado. De esos grafos, se analizó el ritmo y rango de valores a lo largo de la ruta de interés (recompensa de nivel máxima), para las características seleccionadas (de las ocho posibles se seleccionaron seis, la recompensa de nivel, motivación de nivel, riesgo, motivación de bonus, recompensa de bonus y distancia, dado que en [24] demostraron ser las que mayor relación tenían con la etiqueta de experiencia de juego).

Después de proponer diferentes conjuntos de valores objetivo, se encontraron aquellos que se muestran en la Tabla 1. En general, los valores de ritmo (μ), variación del ritmo (σ), valor máximo (λ) y valor mínimo (τ) fueron definidos con base en el valor más frecuente o en el promedio de las medidas observadas en los niveles evaluados.

Por otro lado, los pesos de la exactitud de ritmo (α), exactitud de rango de valores (β) y el peso (w) fueron propuestos con base en la importancia Gini de cada medida (la importancia Gini es una medida de relevancia que se usa dentro del método de selección de características ExtraTreesClassifier, el cual se usa como parte del método de evaluación de niveles seleccionado).

Si una característica tiene mayor importancia Gini que otra, entonces su peso (w) debe ser mayor en la medida de similitud. Para determinar esta relación, sumamos las importancias Gini de cada característica seleccionada y calculamos la aportación proporcional de cada una de ellas para obtener cada peso (w_i).

Si el ritmo de una característica tiene una mayor importancia Gini que su valor máximo o su valor mínimo, entonces la ponderación de la exactitud de ritmo (α) debe ser mayor que la de la exactitud de rango de valores (β) y vice versa.

Para ello, se experimentó con varios valores de α y β , considerando las importancias Gini y buscando una mejor evaluación en la medida de similitud en los niveles ejemplo.

Cabe destacar que, como parte de esta búsqueda de los valores objetivo, se identificó que la *motivación de bonus*, en el caso de la experiencia de juego *simple* causaba una mala identificación de las experiencias de juego, reduciendo la similitud calculada para los niveles que sí correspondían a dicha experiencia; por lo tanto, para esa experiencia de juego no se consideró a la *motivación de bonus* como una característica a evaluar dentro de la medida de similitud.

4.1. Evaluación usando la medida de similitud propuesta

Con base en los valores objetivo definidos en la Tabla 1 los niveles creados obtuvieron las evaluaciones mostradas en la Tabla 2, la cual, muestra que cada nivel tienen un valor de similitud cercano a 1 cuando se evalúa con la medida de similitud ajustada a su correspondiente experiencia de juego y una evaluación cercana a 0 para las medidas de similitud ajustadas a otras experiencias de juego, por lo tanto, se cumple que la similitud es alta cuando se ajustan los valores objetivo de manera adecuada para cada experiencia de juego.

Esto le da la utilidad a la medida de similitud de ser usada en sistemas de inteligencia artificial dando la capacidad de determinar si una experiencia de juego especificada está dentro de una porción de nivel o no. Algunos ejemplos de dichos sistemas son: asistentes de diseño (le indicarían un estimado de cumplimiento de experiencia de juego al diseñador), generadores de secciones de nivel de una experiencia de juego dada (ayudarían al sistema de generación para indicar cuando una experiencia de juego está presente en el contenido generado), entre otros.

5. Conclusiones

En este artículo presentamos una medida de similitud para experiencias de juego capaz de adecuarse al menos a tres diferentes experiencias de juego. Nuestra medida de similitud es independiente de las características a evaluar y las correspondientes métricas a usar para medirlas, sin embargo, requiere que se obtenga un ritmo (o un dato similar) así como los valores máximo y mínimo de las características, además de algún criterio que ayude a determinar el conjunto de valores objetivo (la variación del ritmo, los tres aspectos ya mencionados y sus correspondientes ponderaciones).

También es una medida de similitud independiente del jugador, pues no requiere ningún dato proveniente de él para estimar la experiencia de juego presente en una rebanada de nivel de videojuego.

Finalmente, sería interesante implementar esta medida de similitud para otros video juegos, así como para otras experiencias de juego. Al mostrar la utilidad de esta medida de similitud para determinar que otras experiencias de juego están presentes, en el caso de "Pingu Run" o en otro video juego, podría corroborarse la relevancia de la propuesta para el área de la generación procedimental de contenido.

Tabla 1. Valores objetivo para cada métrica en cada experiencia de juego.

Experiencia de juego	Métrica	μ	σ	λ	τ	α	β	w
Simple	recompensa de nivel	2	3	916	0	0.95	0.05	0.29
	motivación de nivel	0	3	1406	0	0.75	0.25	0.26
	riesgo	2	0.001	491	0	0.65	0.35	0.22
	recompensa de bonus	2	2	265	0	0.85	0.15	0.15
	distancia	4	4	5	1	0	1	0.08
Obstáculos	recompensa de nivel	4	3	800	0	0.95	0.05	0.25
	motivación de nivel	2	3	1384	-70	0.75	0.25	0.22
	riesgo	6	4	243	0	0.65	0.35	0.19
	motivación de bonus	2	0.001	160	-184	0.35	0.65	0.13
	recompensa de bonus	4	4	90	0	0.85	0.15	0.11
	distancia	4	4	11	1	0	1	0.1
Saltos	recompensa de nivel	1	0.001	1000	0	0.95	0.05	0.25
	motivación de nivel	1	0.001	1805	-56	0.75	0.25	0.22
	riesgo	5	2	155	0	0.65	0.35	0.19
	motivación de bonus	1	2	530	-68	0.35	0.65	0.13
	recompensa de bonus	5	3	350	0	0.85	0.15	0.11
	distancia	4	4	8	1	0	1	0.1

6. Contribuciones por autor

Conceptualización y metodología, M.A.M.-A., J.A.T.-L.C.A.D. y H.C.; investigación y recursos, M.A.M.-A., H.C., C.A.D y J.A.T.-L.; software, visualización y curación de datos, J.A.T.-L.; validación, H.C. y M.A.M.-A.; análisis formal, M.A.M.-

Tabla 2. Similitud medida en los niveles ejemplo usando los valores objetivo de la Tabla 1.

Nivel	Medida de similitud		
	Obstáculos	Simple	Saltos
obstáculos 1	0.748515	0.216647	0.169015
obstáculos 2	0.741509	0.33168	0.192525
obstáculos 3	0.695833	0.253077	0.285230
simple 1	0.083789	0.843283	0.08732
simple 2	0.197562	0.765775	0.178005
simple 3	0.204252	0.829947	0.183384
saltos 1	0.189115	0.255675	0.975981
saltos 2	0.158059	0.268668	0.702896
saltos 3	0.181520	0.261360	0.970456

A.; escritura - preparación del manuscrito original, J.A.T.-L., C.A.D.; escritura – revisión y edición M.A.M.-A. y H.C.; supervisión, administración del proyecto y adquisición de financiamiento, M.A.M.-A. Todos los autores han leído y están de acuerdo con la versión a publicar del manuscrito.

7. Financiamiento

Este trabajo fue posible gracias al apoyo del gobierno mexicano (CONACyT) y el Instituto Politécnico Nacional a través de las becas de investigación SIP-2083, SIP-20220553, SIP-20221064 y SIP 20220533 de la SIP-IPN; IPN-COFAA e IPN-EDI.

8. Declaración de la junta de revisión institucional:

No aplicable.

9. Declaración de consentimiento informado:

No aplicable.

10. Declaración de disponibilidad de datos

Los autores están comprometidos a proveer acceso a toda la información necesaria para que los lectores puedan reproducir completamente los resultados obtenidos en este trabajo. Los conjuntos de datos usados tienen disponibilidad pública.

11. Conflictos de intereses

Los autores declaran que no hay conflicto de intereses.

Referencias

1. Togelius, J., Kastbjerg, E., Schedl, D., Yannakakis, G.N.: What is procedural content generation. Mario Borderline. In: PCGames '11: Proceedings of the 2nd International Workshop on Procedural Content Generation in Games, no. 3, pp. 1–6 (2011) doi: 10.1145/2000919.2000922
2. De Kegel, B., Haahr, M.: Procedural puzzle generation: A survey. *IEEE Trans. Games*, vol. 12, no. 1, pp. 21–40 (2020) doi: 10.1109/TG.2019.2917792
3. Risi, S., Togelius, J.: Increasing generality in machine learning through procedural content generation. *Nat Mach Intell*, vol. 2, pp. 428–436 (2020) doi: 10.1038/s42256-020-0208-z
4. Shaker, N., Togelius, J., Nelson, M. J.: *Procedural content generation in games*. Springer. Berlin/Heidelberg, Germany (2016)
5. Summerville, A., Snodgrass, S., Guzdial, M., Holmgård, C., Hoover, A. K., ISaksen, A., Nealen, A., Togelius, J.: Procedural content generation via machine learning (PCGML). *IEEE Transactions on Games*, vol. 10, no. 3, pp. 257–270 (2018) doi: 10.1109/TG.2018.2846639
6. Liu, J., Snodgrass, S., Khalifa, A., Risi, S., Yannakakis, G.N., Togelius, J.: Deep learning for procedural content generation. *Neural Comput & Applic* vol. 33, pp. 19–37 (2021) doi: 10.1007/s00521-020-05383-8
7. Togelius, J., Yannakakis, G. N., Stanley, K. O., Browne, C.: 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) doi: 10.1109/TCIAIG.2011.2148116
8. Yannakakis, G. N., Togelius, J.: *Artificial Intelligence and Games*. Springer. Berlin/Heidelberg, Germany (2018). Available online: gameaibook.org (accessed on 8 April 2022)
9. Alvarez, A., Dahlskog, S., Font, J., Togelius, J.: Empowering quality diversity in dungeon design with interactive constrained map-elites. In: *IEEE Conference on Games (CoG)*, pp. 1–8 (2019) doi: 10.1109/CIG.2019.8848022
10. Desurvire, H., Caplan, M., Toth, J. A.: Using heuristics to evaluate the playability of games. *CHI EA '04: CHI '04 Extended Abstracts on Human Factors in Computing Systems*, pp. 1509–1512 (2004) doi:10.1145/985921.986102
11. Feil, J., Scattergood, M.: *Beginning game level design*. Game development series. Thomson Course Technology, (2005)
12. Nacke, L., Drachen, A., Kuikkaniemi, K., Niesenhaus, J., Korhonen, H. J., Hoogen, W. M., De Kort, Y. A.: Playability and player experience research. In: *Proceedings of Digra'09: Breaking new ground: Innovation in games, play, practice and theory* (2009)
13. Gonzalez-Sánchez, J. L., Gutierrez-Vela, F. L., Montero-Simarro, F., Padilla-Zea, N.: Playability: analysing user experience in video games. *Behaviour & Information Technology*, vol. 31, no. 10, pp. 1033–1054 (2012) doi: 10.1080/0144929X.2012.710648
14. Shaker, Noor, et al.: Fusing visual and behavioral cues for modeling user experience in games. *IEEE transactions on cybernetics*, vol. 43, no 6, pp. 1519–1531 (2013) doi: 10.1109/TCYB.2013.2271738
15. IJsselsteijn, W. A., De Kort, Y. A., Poels, K.: *The game experience questionnaire*. Technische Universiteit Eindhoven. (2013)

16. Calvillo-Gamez, E. H.: On the core elements of the experience of playing video games. Doctoral dissertation, UCL (University College London) (2009)
17. Pedersen, C., Togelius, J., Yannakakis, G. N.: Modeling player experience in Super Mario Bros. In: IEEE Symposium on Computational Intelligence and Games, pp. 132–139 (2009) doi: 10.1109/CIG.2009.5286482
18. Marczak, R., van Vught, J., Schott, G., Nacke, L. E.: Feedback-based gameplay metrics: measuring player experience via automatic visual analysis. In: Proceedings of the 8th Australasian Conference on Interactive Entertainment: no. 6, pp. 1–10 (2012) doi: 10.1145/2336727.2336733
19. Mirza-Babaei, P., Nacke, L. E., Gregory, J., Collins, N., Fitzpatrick, G.: How does it play better? exploring user testing and biometric storyboards in games user research. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1499 (2013) doi:10.1145/2470654.2466200
20. Cowley, B., Charles, D.: Behavlets: A method for practical player modelling using psychology-based player traits and domain specific features. *User Modeling and User-Adapted Interaction*, vol. 26, no. 2, pp. 257–306 (2016)
21. Dahlskog, S., Togelius, J.: Patterns as objectives for level generation. *Design Patterns in Games (DPG)*, Chania, Crete, Greece (2013)
22. Khalifa, A., Green, M. C., Barros, G., Togelius, J.: Intentional computational level design. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 796–803 (2019) doi: 10.1145/3321707.3321849
23. Torres-Léon, J. A.: Pingu Run GitHub Repository. Disponible en: <https://github.com/JAlbertoTorres/Pingu-run>
24. Torres-Léon, J. A.: Generación automática de niveles de videojuegos. Tesis de maestría. Centro de Investigación en Computación, Instituto Politécnico Nacional pp. 94–114 (2021)