

Razonamiento analógico: una herramienta en la creación de narrativas

Edgar Morales-Palafox¹ y Rafael Pérez-y-Pérez²

¹ Posgrado en Ciencias e Ingeniería de la Computación, UNAM, D.F., México
edgar_morales_p@comunidad.unam.mx

² Departamento de Tecnologías de la Información,
División de Ciencias de la Comunicación y Diseño,
UAM Unidad Cuajimalpa, D.F., México
rperez@correo.cua.uam.mx

Resumen El presente trabajo propone un modelo computacional basado en el razonamiento analógico aplicado a un agente computacional generador de historias, con la finalidad de darle una nueva alternativa que permita saber cómo continuar la historia en un momento específico, cuando el conocimiento con el que cuenta no es suficiente para determinar el rumbo que debe seguir la narrativa. El razonamiento analógico se elige porque permite comprender problemas poco conocidos, a través de relaciones de similitud con experiencias almacenadas en nuestra memoria. La evaluación del modelo se hace en dos fases: primero se verifica que el modelo resuelva los bloqueos, resultando con una mejoría de 60%. En la segunda se valida que las historias sigan siendo coherentes e interesantes, con la intervención de la analogía. El modelo resulta benéfico si se utiliza cuando el bloqueo surge al inicio o al final de la creación de la narrativa.

Palabras clave: Razonamiento analógico, analogía, procesos cognitivos, cómputo creativo, creatividad, historias.

1. Introducción

Mexica (Pérez, 1999; 2007) escribe argumentos de historias cortas acerca de los antiguos Mexicanos. Algunas veces, cuando el programa intenta generar un nuevo argumento, se presentan situaciones en las que el sistema no sabe cómo continuar la historia; a esto se le denomina *impasse*. Mexica cuenta actualmente con un método propio (MP) para resolver el problema; sin embargo, algunas veces el MP no encuentra una solución, resultando en una historia que termina abruptamente generando una historia inconclusa. Por consiguiente, se pretende dotar al agente escritor con una alternativa más y con ello reducir el número de *impasses* no resueltos.

De acuerdo con la problemática que se mencionó anteriormente, se buscan alternativas, encontrando que el razonamiento analógico es una de las más adecuadas, por dos razones:

1. La ciencia cognoscitiva ha realizado algunas investigaciones en las que se proponen modelos computacionales que emplean el razonamiento analógico (RA) como

una forma de enfrentar los problemas que son poco conocidos. Algunos de estos modelos se presentan a continuación:

- COPYCAT (Hofstadter & Mitchell, 1995) se enfoca en descubrir analogías, en un contexto compuesto por series de letras del alfabeto, de la “A” a la “Z”.
- ACME (Holyoak & Thagard, 1989) el tipo de analogías que realiza son: Efecto de nivelación de solventes y abrir frascos, Enlaces químicos y estira y afloja, Distribución de las velocidades moleculares y distribución de la riqueza, entre otras.
- ACT-R (Salvucci & Anderson, 2001) es una propuesta de arquitectura cognoscitiva: una teoría acerca de cómo funciona los procesos del pensamiento humano, los investigadores realizaron pruebas del razonamiento analógico que se produce entre el sistema solar y el modelo atómico de Rutherford.

Cabe mencionar que los modelos RA no han sido utilizados como parte de la generación de narrativas, lo que hace atractivo este proyecto.

La segunda razón es porque autores como Mitchell (1993), Koestler (1964), Perkins (1981), Hofstadter (1995), Holyoak y Hummel (2002), Dartnall(2002), sugieren la existencia de una relación entre los procesos creativos y la analogía.

Como base las investigaciones anteriores, el objetivo del presente trabajo es crear un modelo computacional que utilice la analogía como una alternativa que permita minimizar el problema del *impasse*. Para cumplir con dicho objetivo, el presente trabajo se dividió en dos fases: en la primera, se trata de verificar que el modelo de RA cumpla con el propósito de minimizar el número de *impasses* que provocan historias inconclusas; en la segunda, dado que el modelo considera proponer una acción para continuar la historia interviniendo en el desarrollo del relato, se evalúa que la narrativa siga manteniendo su alto nivel de coherencia e interés. Para evaluar las historias en el presente trabajo se consideraron las siguientes variables:

- Fluidez: la secuencia de acciones se lleva de una forma clara, entendible para el lector.
- Estructura: el orden y la combinación de las acciones es adecuada.
- Contenido: se verifica el tipo de acciones que son utilizadas en la narrativa.
- Suspense: la variable que indica el grado de expectativa que se genera en el lector.
- Calidad: se verifica el cumplimiento de los criterios establecidos: fluidez, estructura, contenido y suspense.
- Gusto: se mide la satisfacción que le produce el relato al lector.

Para establecer la forma en la que se implementará el modelo en Mexica, es necesario revisar cómo está conformado.

2. Mexica

El sistema escribe historias que son desarrolladas en dos etapas: la primera corresponde a la construcción de la base del conocimiento y la segunda está encargada de construir las historias.

En la primera etapa, el programa obtiene el conocimiento de dos archivos de texto generados por el usuario:

- Diccionario de acciones: se encuentra aquí el conjunto de acciones (odiar, atacar, curar, entre otras) que puede llevar a cabo un personaje. Se asocia con ellas un conjunto de precondiciones, que deben cumplirse para ser utilizadas; por ejemplo, para emplear la acción ‘curar’, el personaje debe estar herido; además, tiene un conjunto de consecuencias o post condiciones que modifican la trama de la historia.
- Las historias previas: representan narrativas hechas por personas, validadas por el autor del sistema.

De ambos archivos se obtiene el conocimiento que será representado a través de estructuras que podrán ser utilizadas como referencia para generar los nuevos relatos. El conjunto de estructuras de conocimiento es llamado conocimiento previo.

En la segunda etapa, con base en el conocimiento previo, el sistema escribe las historias. Para ello, Mexica utiliza dos procesos: *Engagement* y *Reflection*. Durante el estado de *Engagement*, el sistema genera una secuencia de acciones (por ejemplo, la princesa está enamorada del caballero jaguar, el caballero jaguar ama a la virgen, la princesa odia a la virgen). El estado de *Reflection*, evalúa el material generado hasta el momento, y si es necesario lo modifica; además, trata de resolver los *impasses* que surgen durante el desarrollo de la narrativa. Durante la creación de la historia los personajes interactúan, estableciendo ligas emocionales y tensiones, mismas que se describen a continuación.

2.1. Ligas emocionales

Las ligas emocionales se utilizan como precondiciones y post condiciones de las acciones. El sistema hace uso de dos tipos de ellas, principalmente caracterizadas por el tipo de amor que se expresa: el tipo uno representa el amor de hermanos y el tipo dos el amor de pareja. La representación gráfica de las ligas emocionales es a través de flechas. Para el tipo uno la flecha es continua y para el tipo dos la flecha es punteada. Para facilitar el uso de las ligas emocionales dentro del programa, se ha generado una escala donde las emociones irán de -3 a +3. Hay que considerar que los números negativos expresan la emoción negativa (‘odio’), desde poco intenso (-1) hasta muy intenso (-3); mientras que los números positivos expresan a la emoción positiva (‘amor’), desde poco intenso (+1) hasta muy intenso (+3).

A continuación se presentan algunos ejemplos gráficos de ligas emocionales que surgen al momento de ejecutarse una acción.

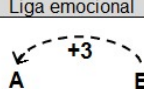
Princesa	A	
Caballero águila	B	
Acción	Liga emocional	Descripción
La princesa está enamorada del caballero águila		Emoción de tipo 2 con intensidad +3
El caballero águila está celoso de la princesa		Emoción de tipo 1 con intensidad -2

Fig. 1. Representación gráfica de algunas ligas emocionales.

2.2. Tensiones

Otro elemento de gran relevancia en la estructura de las historias son las tensiones; por ejemplo, cuando la vida o salud de algún personaje está en riesgo, cuando dos personajes están enamorados de un tercer personaje, cuando muere un personaje, entre otras. Al igual que las ligas emocionales, las tensiones surgen como consecuencia de ejecutar alguna acción; sin embargo, no todas las acciones las disparan. La representación gráfica de las tensiones se hace a través de flechas acerradas, algunas tensiones se presentan en la Figura 2.

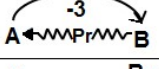

Príncipe	A	
Enemigo	B	
Tlatoani	C	
Acción	Representación	Descripción
El enemigo rapta al príncipe		Emoción de tipo 1 con intensidad -3 y tensión de prisionero
El Tlatoani hiere al enemigo		Emoción de tipo 1 con intensidad -3 y tensión de salud en riesgo

Fig. 2. Representación gráfica de algunas tensiones.

Hasta aquí se han revisado en forma general los elementos que conforman el sistema Mexica, ahora es necesario presentar cómo está conformado el razonamiento analógico, para después plantear el modelo que hace uso de la analogía al momento de surgir un *impasse*.

3. Razonamiento analógico

Siguiendo a Gentner (1983), Lawson (1993), Newton y Newton (1995), se define que la analogía es: la comparación entre dos dominios: uno más conocido y otro poco conocido, que comparten información de tipo relacional.

Los autores Hall (1989), Kedar-Cabelli (1988), Holyoak et al. (1989), proponen que el razonamiento analógico puede ser dividido en tres procesos:

- **Recuperación** (o selección) de una representación del análogo más conocido, de entre un número de posibilidades disponibles en una base de conocimientos, dada una representación del análogo poco conocido.
- **Mapeo** entre los elementos que representan el análogo más conocido y los elementos que representan el análogo poco conocido.
- **Transferencia** desde el análogo más conocido al análogo poco conocido, de la estructura y el significado del contenido.

Expuesto el panorama general de lo que es el RA, se presenta cómo está conformado el modelo.

4. Modelo computacional de RA

El modelo comprende dos etapas: identificación de dominios y uso de analogía. La base del conocimiento que contiene las historias creadas por personas representa el dominio más conocido. En cambio, la narrativa que está en proceso de escritura representa el dominio menos conocido. Ambos están establecidos en el ámbito de las narrativas.

Una vez establecidos los dominios, es necesario identificar los elementos que relacionará la analogía. Para ello, se consideran los siguientes puntos:

- Las historias están formadas por una secuencia de acciones.
- Cada acción puede tener asociada un conjunto de emociones o tensiones, ninguna o ambas.
- Tomar en cuenta el momento en el que surge el *impasse* al generar la historia.

Con base en los puntos anteriores, los elementos a relacionar del dominio poco conocido, estarán formados por las emociones y tensiones que se disparan como resultado de ejecutar la acción en el momento que surge el *impasse*. El dominio más conocido se obtiene de una forma similar al del dominio poco conocido, sólo que se construye a partir de las emociones y tensiones que surgen de las acciones que se encuentran en las historias previas.

Si para resolver el problema sólo se considera la última acción en la que surge el *impasse*, y ésta se busca en el dominio más conocido (conjunto de historias previas), únicamente se obtendría una analogía entre acciones; la intención es ir más allá, es decir, saber qué produjo la última acción, obtener más información. Por ejemplo, la muerte de un personaje puede originarse por lesiones de un accidente, por otro personaje que lo atacó o porque fue sacrificado. Por esta razón, se considera adecuado tomar las últimas dos acciones consecutivas, generando pequeños casos concretos que ayudan a tener una visión más amplia de lo que está ocurriendo en la narrativa y con ello obtener una analogía más adecuada.

Los casos contendrán la secuencia de dos acciones representadas en términos de emociones y tensiones. Para ello, se genera una nueva estructura, llamada Elemento de Transición (ET). Esta estructura contiene las ligas emocionales y tensiones que surgen al pasar de una acción a otra de manera consecutiva. Por ejemplo, si una historia previa contiene esta secuencia de acciones: la virgen curó al caballero jaguar y la virgen se enamoró del caballero jaguar, en la primera acción se dispara una emoción de amor de hermanos intenso, que representa el agradecimiento del caballero jaguar hacia la virgen por haberlo curado, y estas ligas emocionales son las que se guardan en el ET; en la segunda acción surge una emoción de amor de pareja intenso; al igual que la acción anterior, las emociones se guardan en otro ET, que es consecutivo al anterior. Por otra parte, también es creada una estructura que agrupa los ET, con la finalidad de tener identificada la forma en que cambian las emociones y tensiones a través del tiempo en toda la historia previa. Esta estructura se denomina Secuencia de Transformación (ST). La Figura 3 presenta un ejemplo de emociones y tensiones que surgen como resultado de la interacción entre los personajes; es decir, cada personaje adquiere el conocimiento de lo que ocurre a su alrededor en términos de emociones y tensiones al ejecutarse cada acción, Esto es con la finalidad de enseñar cómo se obtienen los elementos que conforman cada ET y asimismo la construcción del ST.

virgen	A		
caballero jaguar	B		
Acción	emociones y tensiones		Estructuras de conocimiento ST
	virgen	caballero jaguar	
1 La virgen CURÓ al caballero jaguar			
2 La virgen SE ENAMORÓ del caballero jaguar			

Fig. 3. Representación de las estructuras de conocimiento.

Teniendo establecidos los dominios y sus elementos, únicamente falta utilizar la analogía; por ello, se establecen los tres procesos del razonamiento analógico:

1. En la recuperación el objetivo es seleccionar del dominio más conocido dos estructuras consecutivas que sean análogas en cuanto a emociones y tensiones, con las últimas dos estructuras que se generaron al momento de surgir el *impasse*, las cuales representan el dominio poco conocido.
2. El mapeo consiste en relacionar las emociones y tensiones que existen en las estructuras ST que conforman ambos dominios, más conocido y poco conocido; por consiguiente, los elementos de ambos dominios se agrupan de tres formas:
 - (a) Elementos de relación presentes en ambos dominios.
 - (b) Los elementos del dominio poco conocido que no están presentes en el dominio más conocido.
 - (c) Los elementos del dominio conocido que no están presentes en el dominio poco conocido.

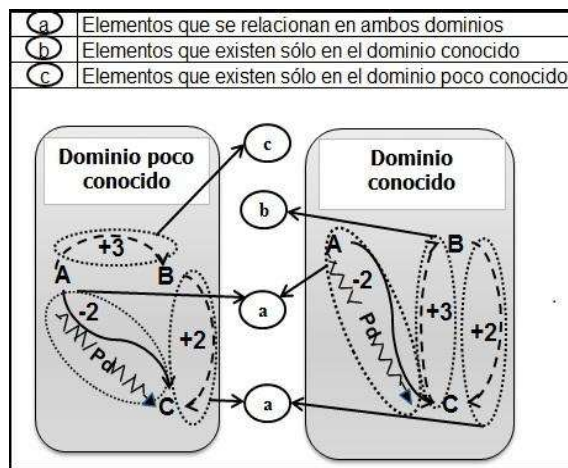


Fig. 4. Mapeo entre dominios.

El mejor caso es cuando todos los elementos quedan en el primer grupo, haciendo una analogía perfecta, pero no siempre es así, habrá casos en donde los elementos estén en las últimas dos agrupaciones, es decir sin relación; por lo que, se evalúa qué tan fuerte es la analogía para tomar una decisión en caso de tener varias alternativas. Por esta razón, se establece una tabla con los posibles elementos que pueden quedar sin relación asignándoles un valor en porcentaje que será restado al cien por ciento de una analogía perfecta ver Tabla 1.

Tabla 1. Valores para los elementos no relacionados.

Elemento sin relación	Porcentaje
Tensiones disparadas por postcondición	20
Tensiones que dispara Mexica	15
Emociones	10
Emociones derivadas por personajes vinculados	5

- La transición se da una vez que se elige la analogía y se procede a determinar cuál es la siguiente acción con la que se debe continuar la historia. Para esto, se establece una relación entre los ET y las estructuras de conocimiento propias de Mexica, ya que en estas estructuras se encuentran las posibles siguientes acciones.

5. Implementación del modelo

En módulo de *Reflection* es donde el modelo de RA se implementa, teniendo como resultado dos alternativas para enfrentar el problema del *impasse*. A continuación se presenta un fragmento de historia donde surge el *impasse*, y es resuelto a través del RA. La creación de la narrativa inicia con la acción, ‘la virgen curó al caballero Jaguar’, en ese momento el sistema dispara una emoción intensa de amor de hermanos, representando el agradecimiento por haberlo curado. En la segunda acción, Mexica decide utilizar la acción, ‘el caballero jaguar asalta a la virgen’; ésta dispara un odio intenso de la virgen hacia el caballero jaguar junto con una tensión de posible peligro; también surge un odio del caballero jaguar hacia sí mismo, por asaltar a quien lo curó. En este momento al no encontrar en su conocimiento alguna acción que le permita continuar la narrativa, surge el *impasse* (ver Figura 5).

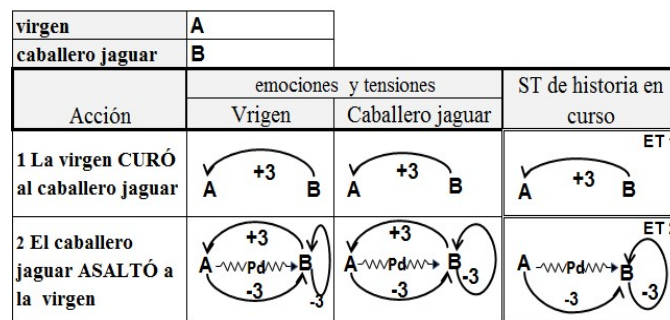


Fig. 5. Surgimiento de *impasse*.

El modelo RA busca la estructura análoga que pase de un momento de amor fraternal a un momento de odio. En este caso se encuentran dos posibles estructuras análogas en el dominio más conocido:

- La primera solución se encuentra en la Secuencia de Transformación dos, específicamente en los ET 5 y ET 7. Las estructuras cambian de un amor fraternal a un odio (relación 1 y 2). La analogía se muestra en la Figura 6.

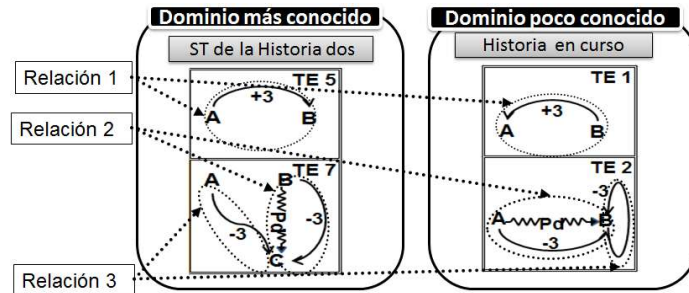


Fig. 6. Primera situación análoga.

- La segunda solución encontrada está en la Secuencia de Transformación tres, en los ET 3 y ET 5. Esta estructura de igual forma es análoga. Es importante mencionar que en esta analogía hay dos personajes ligados con el personaje afectado; como ambos cumplen la misma función, se toma como si fuera uno solo (relación 3). Además, en el dominio origen sobra la tensión Pr; por consiguiente, en la evaluación tiene menor grado de similitud; dicha solución se muestra en la Figura 7.

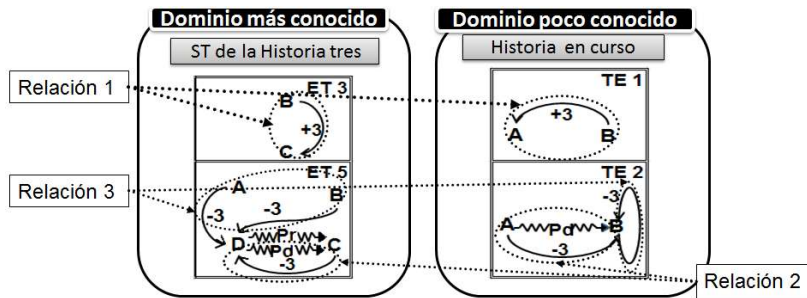


Fig. 7. Segunda situación análoga.

De acuerdo con la agrupación de elementos que conforman los dominios y la tabla de elementos sin relación presentada anteriormente, la primera analogía tiene todos sus elementos relacionados; por lo tanto, tiene un grado de similitud de 100%, mientras que la segunda tiene una tensión de prisionero sin relacionar; por consiguiente, al 100% se le resta el 20% correspondiente al valor de la tensión sin relación, dando como resultado un grado de similitud de 80 %; por lo tanto, la primera analogía es más adecuada para solucionar el *impasse*. Ahora sólo falta determinar cuál es la siguiente acción. Se busca, para ello, el vínculo que hay entre el ET y las estructuras

propias de Mexica. Es en estas últimas donde se encuentran las siguientes posibles acciones; en este caso, corresponde a una estructura con una tensión y tres ligas emocionales, que propone como siguiente acción ‘atacar’, por lo que se continua la narrativa escribiendo: ‘la virgen atacó al caballero jaguar’.

6. Evaluación del modelo

La evaluación del modelo de razonamiento analógico es realizada en dos fases: la primera consiste en verificar que el modelo de RA cumpla con el propósito de minimizar el número de *impasses* que provocan historias inconclusas. Para cumplir con dicho objetivo, se generaron en el sistema Mexica diez casos con *impasse*, y para resolverlos se utilizaron ambas alternativas: método propio y razonamiento analógico. Por un lado, el método propio no tuvo éxito dado que no pudo resolver ninguno de los *impasses*, por otro lado al hacer uso de RA se pudieron resolver seis de los diez casos; con ello se pudo comprobar que el RA es una herramienta apta para la solución del *impasse*.

La segunda fase consistió en verificar que la acción que resuelve el *impasse* propuesta por el RA, fuera tan coherente e interesante como la que se genera a través del método propio de Mexica.

Para evaluar esta fase se utilizó una muestra conformada por 30 sujetos, 13 mujeres y 17 hombres de entre 19 y 25 años ($M = 21$), que fueron seleccionados de forma no probabilística, en una universidad pública al norte del Distrito Federal. Se utilizó un instrumento que fue desarrollado para el presente estudio, que constó de seis reactivos tipo viñetas con formato de respuesta tipo *Likert* de cinco puntos, en los que se contemplaron seis elementos primordiales de la narrativa: fluidez, estructura, contenido, suspenso, calidad y gusto, con el objetivo de determinar, a través de ellos, la coherencia y el interés de las historias generadas por el RA. Además, es considerado el momento (al principio, a la mitad y al final) en que aparece el *impasse* al generar la narrativa.

A través de los resultados se observó que los sujetos tenían diferentes opiniones de las alternativas según el momento en el que aparece el *impasse*. Por un lado, les gustan más las historias resueltas con el modelo de razonamiento cuando surge el *impasse* al principio y al final (ver Figura 8); sin embargo, cuando surge a la mitad de la narrativa el relato gusta más utilizando el MP (ver Figura 9). Por otro lado, señalan que las historias generadas utilizando ambas alternativas mantienen la coherencia e interés. No obstante al observar los indicadores en conjunto existen ciertas diferencias. Por ello, se realizó un análisis estadístico a través de la prueba *t* de Student, resultando en que no hay diferencias estadísticamente significativas al iniciar ($t(29) = .914$, $p > 0.05$), a la mitad de la historia ($t(29) = .868$, $p > 0.05$), ni al final ($t(29) = .029$, $p > 0.05$); dicho de otra forma, ambas herramientas son igual de efectivas para resolver el problema del *impasse*; sin embargo, si el análisis se hace por cada variable, la estructura es la única variable que tiene una diferencia significativa cuando el *impasse* surge al final de la narrativa (ver Figura 8).

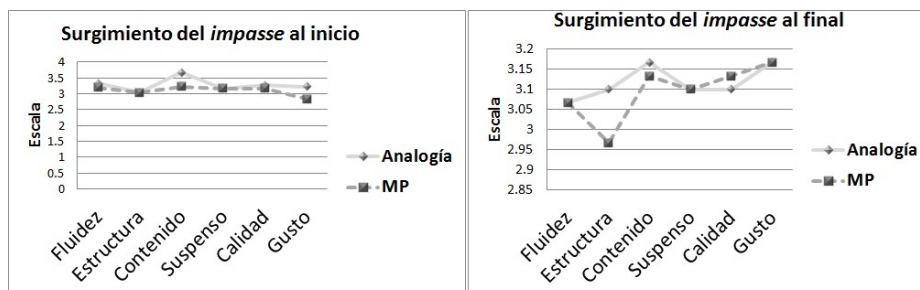


Fig. 8. Representación gráfica del surgimiento del *impasse* al inicio y al final.

En la gráfica de la Figura 9 se observó que los sujetos de estudio aprecian que hay diferencia en la variable de estructura. Utilizando la analogía mantiene un buen grado de estructura; sin embargo, disminuye el grado haciendo uso del método propio de Mexica. Probablemente esto se deba a que son pocas las estructuras de conocimiento con las que cuenta el sistema para finalizar la historia.



Fig. 9. Representación gráfica del surgimiento del *impasse* a la mitad.

7. Conclusiones

El objetivo de la presente investigación ha sido crear un modelo computacional basado en el razonamiento analógico, que pudiera aplicarse en los sistemas generadores de historias como una alternativa que apoyara el desarrollo del relato, con lo cual se pudo concluir que sí es factible crear este tipo de modelos. El modelo de razonamiento analógico logra resolver seis de cada diez casos de *impasse* que el método propio de Mexica no resuelve, resultando benéfico ya que reduce a 40% el número de historias inconclusas. Para reducir más el porcentaje de historias inconclusas es necesario hacer más re combinaciones de estructuras de conocimiento; en el presente proyecto sólo se crearon casos con dos estructuras consecutivas, pudiéndose tomar tres, o más estructuras consecutivas para generar la analogía.

Las variables: fluidez, estructura, contenido, suspense, calidad y gusto, permitieron hacer una adecuada evaluación de las historias en cuanto a coherencia e interés. Los resultados de la evaluación sugieren que Mexica puede utilizar el RA, cuando el *im-*

passee surge al inicio o al final de la narrativa, dado que en todos los casos la evaluación realizada al modelo de RA fue igual o superior a la del método propio de Mexica. A través de los resultados estadísticos se concluye que el modelo de razonamiento analógico es adecuado como una alternativa que apoya en la generación de narrativas. Mexica ahora es uno de los primeros en explorar la creación de narrativas utilizando un modelo de razonamiento analógico, mostrando ser apto en los casos que logra resolver. Por último, el modelo ayuda al entendimiento de la analogía a través de procesos cognoscitivos.

Referencias

1. Dartnall, T.: Creativity, Cognition and Knowledge: An Interaction. London: Praeger (2002)
2. Gentner, D.: Structure-Mapping: a Theoretical Framework for Analogy. Cognitive Science. The MIT Press. Vol. 7, pp. 155-170 (1983)
3. Hall, R.: Computational approaches to analogical reasoning: A comparative analysis. Artificial Intelligence, 39-120 (1989)
4. Hofstater, D., & Mitchell, M.: The copycat project: A model mental fluidity and analogy-making. (D.a. In Hofstadter, Ed.) Basic Books. (1995)
5. Hofstater, D.: Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought. Basic Books (1995)
6. Holyoak, K. & Hummel, J.: Understanding analogy within a biological symbol system. In D. Gentner, K.J. Holyoak and B.N. Kokinov (Eds.), The Analogical Mind: Perspectives from Cognitive Science. London: The MIT Press, 161-195 (2002)
7. Holyoak, K. J. & Thagard P.: Analogical mapping by constrain satisfaction. Cognitive Science. 13 (1989)
8. Kedar-Cabelli, S.: Analogy - From a unified perspective. Analogical Reasoning, ed. D.A. Helman, Kluwer Academic Publishers. Dordrecht, 65-103 (1988)
9. Koestler, A.: The Act of Creation. London: Hutchinson. (1964)
10. Lawson, A. E.: The Importance of Analogy: A Prelude to the Special Issue. Journal of Research in Science Teaching. 30 (10), 1213-1214 (1993)
11. Mitchell, M.: Analogy Making as Perception: A Computer Model. Cambridge, Mass.: Bradford Books. (1993)
12. Newton, D. & Newton, L. Using Analogy to Help Young Children Understand. Educational Studies, Vol. 21 (3), 379-393 (1995)
13. Pérez, R.: MEXICA: a Computer Model of Creativity in Writing. PhD dissertation. University of Sussex (1999)
14. Pérez, R.: Employing Emotions to Drive Plot Generation in a Computer-Based Storyteller. Cognitive Systems Research. Vol. 8, number 2, 89-109 (2007)
15. Perkins, D.: The Mind's Best Work. Cambridge, Mass.: Harvard University Press. (1981).
16. Salvucci, D. D., & Anderson, J. R.: Integrating analogical mapping and general problem solving: The path-mapping theory. Cognitive Science 25, 67-110 (2001)