

LA EVALUACIÓN EN INTELIGENCIA ARTIFICIAL

*Nelson R. Becerra Correa**

*Deicy Alvarado Nieto***

En este artículo se hace un breve análisis sobre los procesos de evaluación en Inteligencia Artificial. Se plantea la forma como ésta se presenta en sistemas expertos, máquinas de aprendizaje, redes neuronales y algoritmos genéticos.

Como fase final se analiza su utilización en el marco de la Teoría de Juegos, presentando una función de evaluación estática para el juego de Mancala.

Evaluación

El término evaluación puede interpretarse de diversas maneras dependiendo del área en que se utilice. La definición que aparece en el DRAE (Diccionario de la Real Academia Española) para el término evaluar es: “señalar, estimar, apreciar, calcular el valor de una cosa”. García Ramos (1989)¹ propone que “la evaluación es un proceso que busca información para la valoración inmediata y la toma de decisiones sobre el valor de algo”.

Cuando se habla de evaluación, ésta inmediatamente se asocia con aspectos educativos. A través de la historia este proceso se ha realizado con el fin de medir lo que el estudiante ha aprendido². Sin embargo, la evaluación se evidencia en muchos otros aspectos tales como el empresarial, en donde la “evaluación” es considerada como el nivel de precisión con el que se alcanzan los objetivos; a nivel gubernamental se habla de la medida en que se alcanzan las metas planteadas. Existen programas de evaluación de la violencia, evaluación farmacéutica y cosmética de los productos; también se habla de evaluación de habilidades en psicología y algunas terapias, evaluación de estándares, etc.

Desde el punto de vista tecnológico puede hablarse de evaluación del nivel de rendimiento del hardware, o de la robustez, exactitud, eficiencia y eficacia de los programas de software. En el campo de la inteligencia artificial es posible hablar de la evaluación de sistemas inteligentes; de igual forma se cuenta con herramientas inteligentes tales como las funciones de evaluación, las redes neuronales, los algoritmos genéticos, los sistemas expertos y las máquinas de aprendizaje automático, que contribuyen en la realización de diversos tipos de evaluación.

* Ingeniero de Sistemas, Magister en Ingeniería de Sistemas Universidad Nacional de Colombia, profesor Universidad Distrital F.J.C., adscrito a la Facultad Tecnológica

** Ingeniera de Sistemas y Especialista en Multimedia Educativa Universidad Nacional de Colombia, aspirante a Doctor en Ciencias de la Computación e Inteligencia Artificial Universidad de Oviedo, profesor Universidad Distrital F.J.C., adscrita a la Facultad de Ingeniería

¹ García Ramos, J. M.: Bases Pedagógicas de la Evaluación: Guía Práctica. Ed. Síntesis. Barcelona 1989

² La metodología para la evaluación, es un tema bastante controvertido, ver por ejemplo Evaluación Integral por procesos Cayetano E. Solano [8].

Según Rolston (1995)³ los sistemas expertos “simulan el comportamiento humano mediante la aplicación específica de conocimientos y de inferencias”, es decir, se toman ciertos hechos que luego son *evaluados* frente a un conjunto de reglas. De esta manera puede decirse que el procesamiento que lleva a cabo un sistema experto es en sí un tipo de evaluación. Nebendahl (1988)⁴ afirma que “los sistemas expertos ofrecen ayuda para diagnosticar fallas con mayor rapidez”; por esto pueden ser aplicados en áreas tales como la evaluación para la concesión de créditos en bancos, análisis de siniestros para compañías de seguros, motivos de fallo de las máquinas en la industria, diagnósticos médicos, entre otros. Todos ellos son sistemas evaluadores que permiten obtener información para tomar una decisión específica.

Por otra parte, todos los sistemas de aprendizaje automático (Winston 1992)⁵ presentan tres componentes principales. Las indicaciones no estructuradas sobre lo que se debe hacer, un algoritmo de aprendizaje y un resolvidor de problemas. Según Bahamonde 1997⁶ “el resolvidor de problemas realiza un proceso de generalización o inducción desde los datos de partida”. Se observa entonces, al igual que en los sistemas expertos, que el procesamiento está basado en evaluaciones. A su vez estos sistemas de aprendizaje automático también tienen muchos campos de aplicación: control de tráfico (evaluación del nivel de carros en una calle), evaluación de sistemas expertos, procesos de evaluación educativa en los cuales se tenga en cuenta la evolución y los conceptos fallidos de cada estudiante en particular.

Las redes neuronales⁷ se utilizan principalmente para la clasificación e identificación. Hay muchos tipos diferentes, cada uno de los cuales tiene aplicaciones diversas; una de ellas la evaluación. A continuación, algunos ejemplos de ello:

- Las redes neuronales artificiales pueden evaluar las probabilidades de formaciones geológicas y petrolíferas.

cas y los riesgos de créditos. También se pueden utilizar en la evaluación del desempeño de los empleados de una empresa; en este caso las entradas de la red podrían ser: puntualidad, eficiencia, eficacia y disponibilidad, y la red podría ser implementada mediante la lógica fuzzy (difusa). Otra aplicación de las redes neuronales en este campo es la evaluación del nivel de tráfico para control de los semáforos, cuyas entradas son: número de vehículos, velocidad de circulación, prioridad de las calles, eventos inesperados, etc.

La evaluación mediante programación evolutiva, más exactamente empleando algoritmos genéticos⁸, se presenta de dos formas: la primera tiene que ver con las áreas en las cuales éstos se aplican; segundo, este concepto se encuentra presente como herramienta del algoritmo en lo que se conoce como función de aptitud, la cual mide la capacidad de un individuo para adaptarse al medio.

Los algoritmos genéticos toman como modelo el proceso de la selección natural, en el cual se elige al más apto; en otras palabras, la misma naturaleza evalúa. Este tipo de algoritmo puede evaluar la función $5/(x+1)$ con el fin de optimizarla determinando su máximo en un intervalo definido. Una de sus aplicaciones como herramienta de evaluación permite medir las condiciones en las cuales se obtiene el máximo rendimiento de un grupo de alumnos en una clase, tomando una función de evaluación estática y optimizándola mediante evaluaciones sucesivas. En 1985 Rendell desarrolló búsquedas de algoritmos genéticos para la evaluación de funciones en juegos.

El propósito del presente artículo es centrarse en el desarrollo de lo que se conoce como función de evaluación.

Función de Evaluación

Las funciones de evaluación producen una estimación de la utilidad esperada de un juego correspon-

³ Rolston, David W. Principios de Inteligencia Artificial y Sistemas Expertos, Ed Mc Graw Hill, 1992

⁴ Nebendahl, Dieter. Sistemas Expertos. Marcombo Editores, 1991

⁵ Los sistemas de aprendizaje automático hacen parte de la Inteligencia artificial y son procesos que pretenden simular el razonamiento partiendo de hechos concretos. En: WINSTON, Henry. Inteligencia Artificial. Addison Wesley, 1992

⁶ Bahamonde R, Antonio. Aprendizaje Automático: Tercer Ciclo en Ciencias de la Computación

⁷ Una forma de computación, inspirada en modelos biológicos

⁸ Un algoritmo genético es una técnica de resolución de problemas inspirada en la evolución de los seres vivos. Estos se utilizan principalmente en problemas de optimización y búsqueda.

diente a una posición determinada. El desempeño de un programa para juegos depende considerablemente de la calidad de su función de evaluación; si esta no es exacta llevará al programa a posiciones que aparentemente son buenas, pero que en realidad son desastrosas. Al generar estas funciones se deben tener en cuenta algunos cuidados⁹ tales como:

- La función de evaluación deberá coincidir con la función de utilidad de los estados terminales
- Su cálculo no deberá ser muy retardado, es decir, se debe establecer un compromiso entre precisión de la función de evaluación y su respectivo costo en tiempo
- La función de evaluación deberá reflejar con precisión las posibilidades de ganar.

Aunque la mayoría de las funciones de evaluación de los programas son lineales¹⁰ recientemente algunas no lineales han tenido éxito.

A continuación se presenta un ejemplo de la aplicabilidad de la función de evaluación en el juego de Mancala¹¹.

Descripción del Juego

El juego de Mancala coloca a dos adversarios frente a un tablero que tiene seis cavidades en cada lado y una cavidad como casa, al lado derecho de cada jugador. Inicialmente hay cuatro fichas en cada cavidad, y los jugadores se turnan para tomarlas desde cualquiera de ellas, distribuyéndolas una a una en las siguientes cavidades, en sentido contrario a las manecillas del reloj. Si la última ficha cae en la primera casilla de la derecha del adversario y en esta hay dos o tres más, se toman estas fichas, se las coloca en su casa y se saca del juego (es decir, las captura); luego se continúa tomando las fichas de las cavidades contiguas, si en estas hay dos o tres fichas. El juego termina cuando no hay fichas suficientes para hacer capturas; gana el jugador que al final del juego tenga más fichas en su casa.

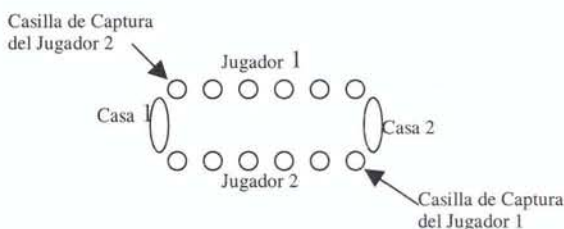


Figura 1. Tablero de Mancala

Análisis de la Función de Evaluación para el Juego De Mancala¹²

El proceso de cálculo de un número que refleje la calidad del tablero se conoce como función de evaluación estática. El procedimiento es realizado por un evaluador estático; el número que se calcula se conoce como resultado de la evaluación estática. A continuación se presenta la función de evaluación estática hallada para el juego de mancala.

$$f(x) = casa_compu - casa_huma + (3(e^{-|n-1|} + (\sum_{j=0}^5 e^{-|(m_{oj}-rp1)(m_{oj}-rp2)|} * e^{-|n-1|}))) - (3(e^{-|n-1|} + (\sum_{j=0}^5 e^{-|(m_{ij}-ro1)(m_{ij}-ro2)|} * e^{-|n-1|})))$$



⁹ Russell, Stuart y Norvick, Peter. Inteligencia Artificial: Un enfoque Moderno. Prentice Hall. 1996

¹⁰ Una función lineal tiene la forma f(x) = ax, donde a es un número real distinto de cero

¹¹ Juego Africano; uno de los juegos de estrategias más antiguos del mundo, aproximadamente de 3500 años de antigüedad

¹² Ver BECERRA C., Nelson. Estrategias de Búsqueda en el Juego de Mancala. Tesis de Maestría Universidad Nacional de Colombia, 1998

En donde:

Casa_compu: representa el número de fichas que posee el computador en su casa.

Casa_huma : número de fichas que posee el adversario humano en su casa.

$e^{1/2n-1/2}$ permite determinar el número de casillas que quedarían en cero una vez distribuidas las fichas.

n: es el número de fichas que hay en la casilla de captura del oponente.

m_{0j} : número de fichas que tiene el adversario en la casilla $0j$, donde j varía de uno a seis

rp1: almacena las fichas necesarias para llegar de la cavidad que en ese momento se analiza (m_{0j}) a la cavidad de captura ($m_{1,6}$). Es de anotar que las fichas sólo alcanzan a dar dos vueltas como máximo

rp2: representa el doble número de fichas necesarias para llegar de la cavidad que se analiza (m_{0j}) a la de captura ($m_{0,0}$).

ro1: son las fichas necesarias para llegar desde la cavidad del oponente ($m_{1,j}$ con J variando de uno a seis) a la cavidad de captura de captura ($m_{1,6}$).

ro2: doble número de fichas para llegar de la cavidad que se analiza ($m_{1,j}$) del oponente a la cavidad de captura ($m_{0,0}$).

A continuación un breve análisis tendiente a obtener el modelo matemático anterior. La función debe tener en cuenta las fichas que hasta ese momento se han capturado por parte del computador; a esto es necesario restarle el número de fichas que el oponente humano ha capturado.



Figura 2: Tablero n luego de una exploración

Si se tomara únicamente la función de evaluación como la relación entre casa-comp (número de fichas capturadas por el computador) – casa humano (número de fichas capturadas por el oponente), el table-

ro anterior tendría un valor óptimo y sería la opción a jugar. En el desarrollo del juego el computador trataría de llegar a ese tablero para ofrecérselo al adversario.

Cuando el adversario distribuya sus fichas, toma la casilla donde se encuentran seis de ellas; esta es una jugada ideal, pues con ello capturaría 12, lo cual sería fatal para el oponente (computador). Por consiguiente, además del número de fichas en cada casa es necesario analizar si en la casilla de captura la primera del lado correspondiente a cada jugador posee una o dos fichas, de tal manera que cualquier jugador, una vez distribuya sus fichas, complete dos o tres indispensables para la captura; también se debe prever si se poseen las fichas necesarias, para que al distribuir las últimas caiga en la casa del oponente y, de esta forma poder capturarlas.

Otro elemento a tener en cuenta es saber qué puede hacer el oponente, si éste también posee el número de fichas suficientes para alcanzar la casilla de captura y si se poseen 2 o 3 fichas en las casillas, fundamentalmente en la de captura, pues este tablero sería malo y debería tener un valor bajo.

El hecho de incluir en el tablero esta última parte del análisis tiene como objetivo revisar, mediante la función de evaluación, un nivel más, y con esto contrarrestar en parte el efecto horizonte¹³. Además, como al analizar el tablero no se sabe a qué jugador le corresponde, se deben tener en cuenta tanto la parte positiva como la parte negativa del tablero.

La función que permite determinar el número de casillas que quedan en cero luego de distribuidas las fichas es $e^{-1/2n-1/2}$, donde "n" es el número en la casilla de captura del oponente; por consiguiente esta función toma su mayor valor cuando n es igual a uno (1). Esto garantizaría que al distribuir las fichas en esa cavidad haya máximo tres, suficientes para hacer la captura.

La función que permite verificar si hay fichas suficientes en cada una de las casillas para acceder a la casilla de captura es en donde la fracción toma

¹³ El efecto horizonte se presenta cuando aparentemente se tiene una jugada fabulosa, pero que en el desarrollo del juego algunas jugadas más adelante resulta fatal.

su máximo valor cuando $rp1$ o $rp2$ sea igual a $m0j$, con el fin de indicar que hay fichas suficientes para colocar la última piedra en el hueco de captura. La sumatoria permite comprobar cada uno de los huecos. La multiplicación por $e^{-1/2n-1/2}$ garantiza que es válida cuando el hueco de captura hay dos o tres.

La componente siguiente, con signo negativo, interpreta el mismo concepto expuesto anteriormente:

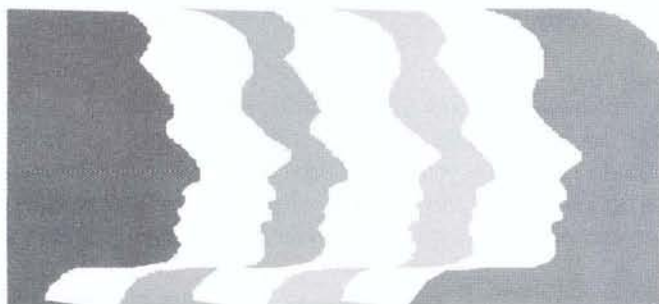
De esta forma queda completa la función de evaluación estática para la versión original de Máncala.

Conclusiones

- En todas las actividades de la vida diaria se llevan a cabo procesos de evaluación, los cuales permi-

ten una retroalimentación necesaria para una mejor adaptación al medio

- Todos los sistemas inteligentes artificiales deben incorporar procesos de evaluación, con el fin de simular en una forma más precisa el razonamiento humano
- Se hace necesario plantear la creación de nuevas metodologías para la incorporación de sistemas evaluativos en Inteligencia Artificial
- Es evidente la necesidad de crear funciones de evaluación que cambien con el transcurso del tiempo para que reemplacen las actuales funciones de evaluación estática.



REFERENCIAS BIBLIOGRÁFICAS

- BAHAMONDE R, Antonio. Aprendizaje Automático: Tercer Ciclo en Ciencias de la Computación. Centro de Inteligencia Artificial Universidad de Oviedo, Gijón. Curso Académico 1996-1997
- BECERRA CORREA Nelson R. Estrategias de Búsqueda en el Juego de Mancala, Tesis de Maestría, Universidad Nacional de Colombia, 1998
- CAYETANO ESTÉVEZ Solano, Evaluación Integral por Procesos. Una Experiencia Construída Desde y en el Aula, Cooperativa, Ed. Magisterio, 1997
- GARCÍA RAMOS, J.M. Bases Pedagógicas de la Evaluación: Guía Práctica, Ed. Síntesis, Barcelona, 1989
- NEBENDAHL, Dieter. Sistemas Expertos, Marcombo Editores, 1991
- ROLSTON, David W. Principios de Inteligencia Artificial y Sistemas Expertos. Ed. Mc. Graw Hill, 1992
- RUSSELL, STUART Y NORVICK, Peter. Inteligencia Artificial: Un Enfoque Moderno. Prentice Hall. 1996
- WINSTON, Henry. Inteligencia Artificial. Addison Wesley. 1992.