

Evaluador inteligente basado en lógica difusa para simuladores de navegación.

Valles Cancela, José Ignacio ¹, Olivas, José A. ², González Prieto, José Antonio ³

¹ Secretaría General del Estado Mayor de la Armada. Sección de Gestión de la Información y Conocimiento. Centro de Inteligencia Artificial de la Armada. jvalc11@fn.mde.es: (JIVC), ² Universidad de Castilla-La Mancha. Departamento de Tecnologías y Sistemas de Información. Escuela Superior de Informática. JoseAngel.Olivas@uclm.es (JAO), ³ Centro Universitario de la Defensa. Escuela Naval Militar. jose.gonzalez@tud.uvigo.es: (JAGP)

Resumen

Este trabajo se centra en el desarrollo de un sistema de evaluación inteligente, que pueda ser integrado en los simuladores de navegación que la Armada emplea en sus centros de formación, con el objetivo de apoyar a los instructores en el proceso de evaluación de los alumnos. Para lograr este objetivo se propone emplear técnicas de lógica difusa y razonamiento aproximado para, a partir de los datos obtenidos de las simulaciones y en base al conocimiento experto de los instructores, desarrollar un evaluador inteligente para el simulador de navegación de la Escuela Naval Militar (ENM).

1. Introducción

La simulación es una herramienta de un gran valor para la enseñanza, sobre todo en ámbitos como la navegación, en donde no siempre es posible realizar prácticas reales en la mar, en diferentes zonas geográficas, con condiciones densidad de tráfico marítimo variables o en entornos con diferentes condiciones de visibilidad o meteorológicas.

Generalmente, el uso de la simulación en la enseñanza requiere la presencia de instructores expertos que estén en todo momento atentos a la ejecución de los ejercicios por parte de los alumnos, con el objetivo de poder valorar su desempeño y ayudarles a identificar los aspectos a mejorar. En cada ejercicio, un alumno está al mando del puente de gobierno del barco y toma las acciones necesarias para llevar el barco a su destino con precisión y seguridad.

De esta forma, surge la necesidad de establecer un mecanismo de evaluación objetivo y automatizado que permita, por una parte, poder obtener una evaluación precisa del desempeño del alumno, y por otra, mitigar los sesgos naturales relacionados con las decisiones tomadas por humanos en diferentes situaciones. Para poder resolver este problema existen diferentes tipos de soluciones, siendo los sistemas basados en lógica difusa los que mejor permiten modelar la información dinámica y variable procedente del entorno marítimo simulado en los escenarios y plasmar de forma sencilla el conocimiento experto de los instructores.

Este tipo de soluciones ya ha sido empleado en situaciones similares, como por ejemplo para evaluar jugadores [1] y [2], productividad de departamentos [3], robots de cirugía

[4] o gestión de situaciones de emergencia [5], entre otras aplicaciones similares.

2. Objetivos de desarrollo

El objetivo principal del desarrollo es la creación de un sistema de ayuda a la decisión basado en lógica difusa y razonamiento aproximado para la evaluación automática del desempeño de los alumnos en simuladores de navegación, bien como complemento a la observación y orientación proporcionada por los instructores, o para permitir a los alumnos llegar a adiestrarse sin un instructor dedicado, recibiendo una realimentación confiable sobre su desempeño, que les ayude a mejorar.

Para alcanzar el objetivo principal descrito, se establecen los siguientes objetivos parciales:

- 1) El sistema debe ser capaz de extraer la información de los registros del simulador tras el ejercicio y modelar el entorno de la simulación en forma de datos, que permitan llevar a cabo el proceso de evaluación.
- 2) Es necesario poder modelar el conocimiento de los instructores por medio de reglas lingüísticas, de forma que sea posible realizar inferencia con lógica difusa sobre los datos extraídos del simulador para obtener las notas de desempeño.
- 3) Debe tenerse en cuenta que la solución propuesta no esté constreñida a una secuencia determinada de eventos a afrontar y reacciones específicas a imitar, sino que permita hacer frente a escenarios y situaciones cambiantes, sin coartar la iniciativa de los instructores durante el ejercicio.
- 4) El sistema de representación de la información procesada debe ser lo suficientemente interpretable como para permitir entender el motivo de la calificación alcanzada y potenciar la capacidad del alumno para aprender de sus aciertos y errores.

El uso de lógica difusa y razonamiento aproximado evita enfoques difícilmente interpretables y permite aprovechar el conocimiento experto existente. Para ello se empleó el denominado enfoque de *Mamdani*, empleando *term sets* (conjunto de términos de una variable lingüística difusa) con etiquetas lingüísticas para definir las variables de entrada y salida, así como una serie de reglas lingüísticas para representar el conocimiento experto [6].

Figura 3. Ejemplo de la idoneidad de los rumbos mantenidos en uno de los tramos de la derrota.

Además, se evaluó la capacidad de los alumnos para llevar a cabo las maniobras de cambio de rumbo de un tramo a otro, considerando el desplazamiento longitudinal y lateral del barco durante la caída, ver Figura 4.



Figura 4. Ejemplo de cambio de rumbo para el siguiente tramo (fuente: www.cruisersforum.com).

A efectos del nivel de exigencia, se debe tener en cuenta si los tramos son de aguas libres o restringidas, obteniéndose las mejores calificaciones en los casos en los que el barco comienza y finaliza la caída lo más próximo posible al eje de la derrota. Para el cálculo de la nota se miden las distancias al eje en determinados instantes anteriores y posteriores a la distancia mínima de paso al waypoint de cambio de tramo, tal y como se muestra en la Figura 5.

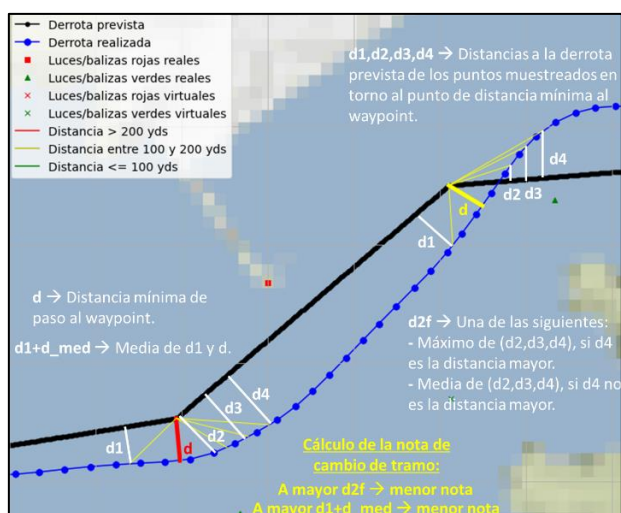


Figura 5. Detalle de las distancias medidas para el cálculo de la nota de caída de rumbo de cambio de tramo.

Para evaluar las distancias de CPA con los contactos, se calcula el CPA con todos los barcos del escenario y se tiene en cuenta si tienen lugar en tramos de aguas libres o aguas restringidas, siendo más exigentes en la distancia de separación en los tramos de aguas libres. Además, en caso de producirse una colisión con otro barco, se añade una penalización en la nota resultante.

Con el objetivo de evaluar la capacidad para navegar por aguas seguras se tiene en cuenta la profundidad por debajo de la quilla, que se registra continuamente durante el ejercicio, así como la aproximación a unas líneas definidas para cada escenario que separan las aguas seguras de las peligrosas, denominadas líneas límite de peligro (LLP). En estos casos, de producirse una profundidad por debajo de la quilla menor a los valores de seguridad establecidos, en

caso de producirse una situación de varada o de un exceso de aproximación o cruce de las LLP en aguas restringidas, se añade una penalización a la nota final resultante.

Para detectar estas situaciones de aproximación excesiva o cruce de las LLP se empleó el algoritmo de agrupamiento DBSCAN, aplicado a los puntos de cada LLP junto con un conjunto de puntos equidistantes muestreados de la derrota realizada, como se muestra en la Figura 6. A mayor porcentaje de puntos de la derrota asignados a los mismos grupos que los de las LLP, mayor será la penalización.

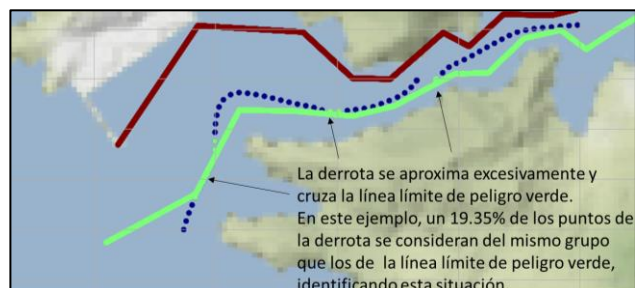


Figura 6. Agrupamiento de la derrota realizada muestreada (puntos discontinuos) y las LLP.

Para realizar el cálculo de las calificaciones parciales se emplea un sistema de inferencia con lógica difusa con el enfoque de *Mandani*. Para ello, se establecen las etiquetas lingüísticas para los datos de entrada y las notas de salida, así como las reglas lingüísticas necesarias para la inferencia. El método empleado para obtener la respuesta definitiva es el denominado del centro de gravedad de la figura de pertenencia resultante en el *term set* de salida [6].

Las Figuras 7, 8, 9 y 10 muestran, a modo de ejemplo, el proceso de cálculo de una de las notas parciales, que evalúa la capacidad para aproximarse y mantenerse en el eje de la derrota en cada tramo, empleando los siguientes *term sets* de entrada y reglas lingüísticas de inferencia difusa.

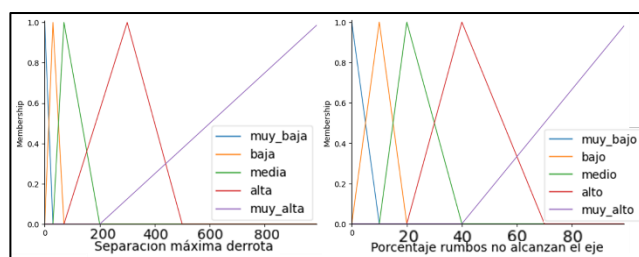


Figura 7. Term sets de los datos de entrada.

Nota A (Aguas restringidas)			
Número	Separación máxima al eje de la derrota	Porcentaje de rumbos que no alcanzan el eje	Nota
1	Muy Baja	-	Muy Alta
2	Muy Baja	&	Muy Alto
3	Baja	-	Alta
4	Baja	&	Muy Alto
5	Media	-	Media
6	Media	&	Muy Alto
7	Alta	-	Baja
8	Alta	&	Muy Alto
9	Muy Alta	-	Muy Baja

Figura 8. Reglas lingüísticas definidas para la inferencia.

La nota se calcula en base a la separación máxima al eje medida en el tramo y el porcentaje de rumbos inadecuados para alcanzar el eje de la derrota. Por ejemplo, con una separación máxima de 60.5 yardas y un porcentaje de rumbos inadecuados de un 83.5%, se generan los siguientes grados de pertenencia a los *term sets* de entrada:

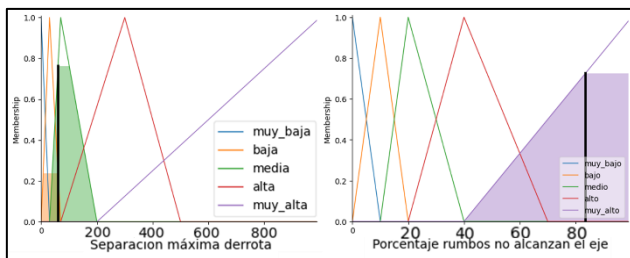


Figura 9. Grados de pertenencia para los datos de entrada.

Tras aplicar la inferencia difusa con las reglas lingüísticas definidas en la Figura 8, se obtiene la figura de pertenencia resultante en el *term set* de la nota de salida de la Figura 10. A continuación, se emplea el método del centro de gravedad para obtener la calificación numérica de salida de 4.65 puntos.

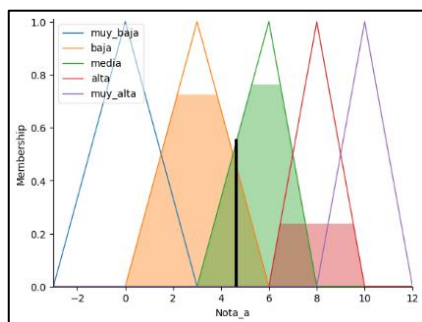


Figura 10. Figura de pertenencia generada en el *term set* de la nota de salida y nota final obtenida.

Finalmente, se calcula la nota consolidada resultante de cada aspecto evaluado, junto con las penalizaciones acumuladas y se obtiene un informe final de calificación en PDF para su almacenamiento en una base de datos y su remisión al alumno por correo electrónico.

5. Resultados

Se realizaron pruebas de los algoritmos KNN y DBSCAN en cuatro escenarios de diferentes dimensiones y características (Ría de Ferrol, Golfo de Cádiz, Aguas costeras de Cartagena y el Estrecho de Gibraltar), obteniéndose el comportamiento esperado y buenos resultados en todos los casos.

Igualmente, se realizaron pruebas de evaluación con tres escenarios diferentes en la ría de Ferrol, cuya calidad de ejecución variaba entre mala, regular y buena. Los resultados de evaluación obtenidos fueron coherentes a las opiniones de los expertos del simulador de navegación y se asignaron las mejores calificaciones al ejercicio de mejor ejecutado, disminuyendo las calificaciones obtenidas en los ejercicios de peor calidad de forma acorde a su ejecución.

En estos escenarios se compaginaron los ejercicios de navegación de precisión con otro tipo de ejercicios como aprovisionamiento en la mar, formaciones navales, etc., sin interferir en el funcionamiento del evaluador, gracias a la posibilidad de definir las zonas de derrota evaluada en el interfaz creado con *Streamlit*. Además, se realizaron reposicionamientos de contactos del entorno para propiciar situaciones de CPA a evaluar o se reposicionó el barco del alumno para cumplir los objetivos de adiestramiento, sin afectar al funcionamiento del evaluador y consiguiendo no coartar la iniciativa de los instructores durante el ejercicio.

6. Conclusiones

El sistema de ayuda a la decisión basado en lógica difusa y razonamiento aproximado desarrollado para la asignación de calificaciones de desempeño de los alumnos en simuladores de navegación permite identificar los aspectos más importantes a evaluar y asigna notas de forma coherente, en relación con el conocimiento experto acumulado y a la ejecución observada de los ejercicios.

Como resultado del proceso de evaluación, se obtienen las calificaciones agrupadas en diferentes áreas de la navegación en base al análisis de los archivos de registro proporcionados por el simulador y se plasman en un informe de desempeño junto con imágenes explicativas, que permiten al alumno apreciar los aciertos y errores cometidos de forma gráfica e intuitiva, así como interpretar el motivo de la asignación de cada calificación.

Referencias

- [1] Romero F, Angulo E, Serrano-Guerrero J, Olivas JA. A Fuzzy Framework to Evaluate Players' Performance in Handball. *International Journal of Computational Intelligence Systems*. 2020; 13(1):549-558
- [2] Li YJ, Ng PHF, Wang HB, Shiu SCK, Li Y. Applying fuzzy integral to performance evaluation in real time strategy game. *International Conference on Machine Learning and Cybernetics*; 2012 15-17 Jul; Xi'an, China. IEEE, 2012. 289 P.
- [3] Ahmed I, Sultana I, Paul S, Azeem A. Employee performance evaluation: A fuzzy approach. *International Journal of Productivity and Performance Management*. 2013; 62:718-735.
- [4] Huang J, Payandeh S, Doris P, Hajshirmohammadi I. Fuzzy classification: towards evaluating performance on a surgical simulator. *Studies in health technology and informatics*. 2005;111:194-200.
- [5] Gul M, Celik E, Gumus AT, Guneri AF. Emergency department performance evaluation by an integrated simulation and interval type-2 fuzzy MCDM-based scenario analysis. *European Journal of Industrial Engineering*. 2016;10(2):196-223.
- [6] Olivas, JA. La Lógica Borrosa y sus aplicaciones. *BOLE.TIC*.2002;24:21-28.
- [7] Olivas, JA. Razonamiento Aproximado. *Materiales del profesor para la asignatura 50689. Máster Universitario en Inteligencia Artificial*. Universidad Internacional de Valencia. 2022.