Búsqueda Neurosimbólica Jerárquica con Reevaluación Selectiva de Transformers para Ajedrez

Álvaro Rodríguez-Gallardo[1,2][0009-0008-2142-2970], Jorge López-Remacho[1,2][0009-0003-7871-3537], Ignacio Aguilera-Martos[1,2][0000-0001-9620-5235], Juan Luis Suárez-Díaz[1,2][0000-0001-8587-4345], Miguel Ángel de la Vega-Rodríguez[2][0009-0002-4910-1163], Alejandro Domingo[2], and Francisco Herrera[1,2][0000-0002-7283-312X]

[1] Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, Granada, España.
[2] Instituto Andaluz Interuniversitario en Ciencia de Datos e Inteligencia Computacional (DaSCI), Universidad de Granada, Granada, España.
arodrgallardo@ugr.es, jorgelerre@ugr.es, nacheteam@ugr.es, jlsuarezdiaz@ugr.es, miguevrgo@gmail.com, adomingo99@hotmail.com, herrera@decsai.ugr.es

Abstract. Los avances recientes en Inteligencia Artificial conexionista han demostrado que los transformers a gran escala pueden alcanzar un rendimiento de nivel maestro en juegos estratégicos como el ajedrez sin recurrir explícitamente a mecanismos de búsqueda. Sin embargo, este enfoque depende de modelos con una gran cantidad de parámetros, lo que incrementa significativamente los costes computacionales y limita la interpretabilidad del proceso de decisión. En este trabajo presentamos Delibera, un marco neurosimbólico que integra transformers en una arquitectura estructurada de búsqueda y reevaluación. Asimismo, proponemos un enfoque jerárquico en el que modelos ligeros generan candidatos en los nodos hoja, mientras que un modelo de mayor capacidad actúa como evaluador secundario en puntos clave del árbol de búsqueda. Los experimentos, respaldados por estudios de tiempos de ejecución sobre diversos conjuntos de puzzles en ajedrez de táctica y estrategia, muestran que la reevaluación compensa eficazmente la menor escala de los modelos base. En particular, la configuración con reevaluación sobre el modelo de 9M alcanza un Elo Blitz de 3075, superando en 180 puntos al transformer 270M empleado de forma aislada. Estos resultados sugieren que la combinación eficiente de modelos conexionistas y simbólicos constituye una alternativa robusta y computacionalmente eficiente al escalado monolítico de modelos para entornos complejos de toma de decisiones.

Keywords: Ajedrez · Transformer · IA Neurosimbólica · Eficiencia · Inteligencia Artificial · Redes Neuronales Profundas.

1 Introducción

Los modelos conexionistas han transformado el campo de la Inteligencia Artificial (IA) en la última década, impulsados por el uso masivo del aprendizaje profundo y, en particular, por la aparición de las arquitecturas transformer [1]. Estos modelos son la base de los grandes modelos de lenguaje (LLMs), con capacidades avanzadas en procesamiento de lenguaje natural, generación de texto y razonamiento multimodal.

El impacto de la IA conexionista se extiende más allá de los LLMs conversacionales, pues los agentes de IA integran razonamiento interno con herramientas externas [2], los dedicados a visión compiten con las redes convolucionales [3] y los transformers a gran escala (LSTs) han demostrado un rendimiento competitivo en juegos estratégicos como el ajedrez [4].

En este dominio, el motor *Searchless Chess* [5] representa un gran avance dado que un transformer a gran escala, de 270 millones de parámetros, entrenado sobre 15000 millones de instancias anotadas por Stockfish 16, alcanza un Elo Blitz estimado de 2895 en Lichess sin emplear ningún mecanismo explícito de búsqueda. Sin embargo, este rendimiento se sustenta en un escalado de parámetros con los costes computacionales asociados y, además, carece de la transparencia propia del razonamiento simbólico.

La IA simbólica [6] sigue siendo indiscutible en entornos de mundo cerrado, donde la búsqueda [7] permite conseguir soluciones de alta calidad. Motores como Stockfish [8] dominan al resto en el juego competitivo, siendo también usado como oráculo para el entrenamiento de LSTs. A diferencia de los modelos conexionistas, los simbólicos ofrecen soluciones explicables, donde bajo la heurística adoptada presenta el mejor resultado alcanzable en el espacio explorado.

La hipótesis central sobre la que se desarrolla el trabajo se centra en que las fortalezas de los modelos conexionistas deben ser explotadas no mediante el escalado bruto de parámetros, sino con una orquestación inteligente dentro de un marco neurosimbólico estructurado. Proponemos que la asignación dinámica de la capacidad computacional, donde se usarán modelos ligeros para la exploración del espacio de estados y modelos de alta capacidad para la reevaluación selectiva, puede superar el rendimiento de arquitecturas monolíticas, siendo este un uso más eficiente de los parámetros que la inferencia estática extremo a extremo.

Con este objetivo, presentamos Delibera, una arquitectura neurosimbólica con dos configuraciones. En primer lugar, una búsqueda básica, DeliberaBasica, que integra un transformer en un árbol de búsqueda clásico. Por otro lado, presentamos DeliberaReev, que consiste en un esquema jerárquico en el que modelos ligeros generan candidatos en los nodos hoja y un modelo de 270 millones de parámetros los reevalúa en puntos clave del árbol. Los experimentos demuestran que este enfoque permite a modelos ligeros superar al transformer de mayor tamaño usado de forma aislada, sugiriendo que la reevaluación de ramas de búsqueda constituye un uso más eficiente de los parámetros que la inferencia monolítica a gran escala.

2 Antecedentes

La IA ha estado históricamente marcada por dos paradigmas: la IA simbólica, basada en la representación explícita del conocimiento e inferencia lógica, y la IA conexionista, basada en modelos sub-simbólicos obtenidos a partir de datos. A continuación, revisaremos sus principios, limitaciones y aplicaciones al ajedrez como banco de pruebas.

2.1 Paradigmas conexionista y simbólico

La IA conexionista, inspirada en las redes neuronales biológicas, evolucionó desde los primeros perceptrones [9] y la retropropagación [10] hasta la revolución del aprendizaje profundo, impulsada por la accesibilidad a grandes conjuntos de datos y recursos computacionales [11]. El culmen fue la arquitectura transformer [1], cuyo mecanismo de atención permitió modelar dependencias de largo alcance con gran eficacia, dando lugar a los LLMs como GPT [12] o LLaMA [13]. Otros desarrollos más recientes como DeepSeek-R1 [14] u OpenAI o3 [15] extienden este paradigma al escalado en tiempo de inferencia con aprendizaje por refuerzo. Sin embargo, los modelos conexionistas presentan limitaciones estructurales relevantes dada la opacidad en su proceso de decisión, la ausencia de razonamiento deliberativo nativo y la dependencia del volumen de datos para generalizar correctamente.

La IA simbólica, fundamentada en la Hipótesis del Sistema de Símbolos Físico [16], opera con la manipulación formal de representaciones explícitas de conocimiento. Sus contribuciones al razonamiento automático [17] y los sistemas expertos [18] tuvieron gran relevancia en las décadas posteriores. En entornos adversariales, el algoritmo minimax [19] y la poda alfa-beta [7] ofrecen un marco formalmente óptimo para la toma de decisiones en juegos de suma cero. A pesar de su potencia, la IA simbólica sufre de cuello de botella en la adquisición de conocimiento, fragilidad en situaciones no previstas y explosión combinatoria del espacio de estados.

La integración neurosimbólica surge precisamente para aprovechar las fortalezas de ambos paradigmas. Trabajos como AlphaGeometry [20] demuestran que combinar reconocimiento de patrones con motores simbólicos permite resolver problemas complejos de forma óptima [21].

2.2 IA aplicada al ajedrez

El ajedrez ha sido históricamente el banco de pruebas por excelencia de la IA [22]. Esto es debido a su complejidad combinatoria y sus reglas bien definidas.

Dominancia simbólica Durante décadas, los mejores motores como Stockfish [8] se apoyaron en una búsqueda alfa-beta muy optimizada y funciones heurísticas. Estas últimas eran codificadas manualmente con conocimiento experto, por lo que eran poco generalizables a escenarios no tenidos en cuenta por el experto.

Irrupción conexionista La introducción de métodos conexionistas demostró que el alto rendimiento de los motores puede emerger del aprendizaje desde los datos y el juego contra sí mismo [23], sin depender de conocimiento explícito. Esto no solo dio lugar a motores competitivos con los simbólicos, sino que transformó la concepción de cómo construir sistemas de decisión en dominios estructurados.

Searchless Chess El trabajo de Ruoss et al. [5] constituye el punto de partida en esta investigación. Los autores entrenan una familia de transformers de 9, 136 y 270 millones de parámetros con aprendizaje supervisado sobre un corpus de, aproximadamente, 10 millones de partidas de Lichess, anotadas con valores de acción calculados por Stockfish 16, generando en torno a 15000 millones de pares tablero-acción. El modelo de 270M alcanza un Elo Blitz de 2895 en Lichess, con nivel de gran maestro, sin usar mecanismos de búsqueda explícita, heurísticas artesanales ni planificación deliberativa. Los tres modelos muestran un comportamiento de escalado monótono, tal que a mayor número de parámetros, mayor cantidad y mejores patrones capturados. Sin embargo, estas políticas puramente *searchless* carecen de mecanismos para la exploración dirigida y la reevaluación dinámica, lo que se manifiesta en indecisión en posiciones de victoria forzada o incapacidad para evitar repetir estados, motivando el enfoque híbrido del trabajo.

Enfoques neurosimbólicos El estado del arte actual en ajedrez está dominado por arquitecturas híbridas. AlphaZero [23] combina la búsqueda mediante árboles de Monte Carlo (MCTS) [24] con una red neuronal que proporciona tanto la política como la función de valor, abandonando el minimax clásico. Por su parte, Stockfish integra una red NNUE [25] exclusivamente como función evaluadora de los nodos hoja, sin que la búsqueda esté guiada por la intuición que puede aportar un modelo conexionista. Ninguno de los dos explora el potencial de un marco híbrido que combine la búsqueda determinista minimax con poda alfa-beta y un transformer para guiar la exploración del árbol, que es precisamente el ámbito que aborda este trabajo.

3 Arquitecturas propuestas

En esta sección presentamos las dos variantes de la arquitectura **Delibera**, un marco neurosimbólico que integra transformers de *Searchless Chess* [5] dentro de un esquema de búsqueda minimax con poda alfa-beta [26]. La versión base, **DeliberaBasica**, incorpora el transformer como heurística de evaluación en los nodos hoja, mientras que la variante, **DeliberaReev**, añade un mecanismo de reevaluación selectiva que reserva el modelo de mayor capacidad para los puntos clave del árbol de búsqueda.

3.1 Arquitectura básica. DeliberaBasica

La arquitectura base usa dos tipos de heurísticas conexionistas: un transformer ligero para la evaluación de nodos hoja y un transformer pesado para guía de la

búsqueda. El algoritmo minimax con poda alfa-beta actúa como esqueleto simbólico, mientras que los transformers *Searchless Chess* [5] se usan como funciones evaluadoras.

Para mejorar la eficiencia, se incorpora una etapa de poda que restringe la exploración a los K mejores candidatos según el transformer pesado, favoreciendo la explotación sobre la exploración. Esta estrategia está respaldada por los autores del modelo base [5], quienes demuestran que la solución correcta aparece sistemáticamente entre los tres primeros movimientos. Además, tras seleccionar los K mejores movimientos, se extiende la búsqueda a W candidatos adicionales, realizando una búsqueda superficial para cada uno. Si su evaluación mejora respecto a la puntuación inicial del transformer pesado, también se explora. Formalmente, dado un estado de tablero s y la lista ordenada de movimientos legales $M(s) = \{m_1, ..., m_N\}$, el algoritmo explora el conjunto

$$\{m_1, ..., m_L\} \quad \text{donde} \quad K \leq L \leq K + W \leq N,$$

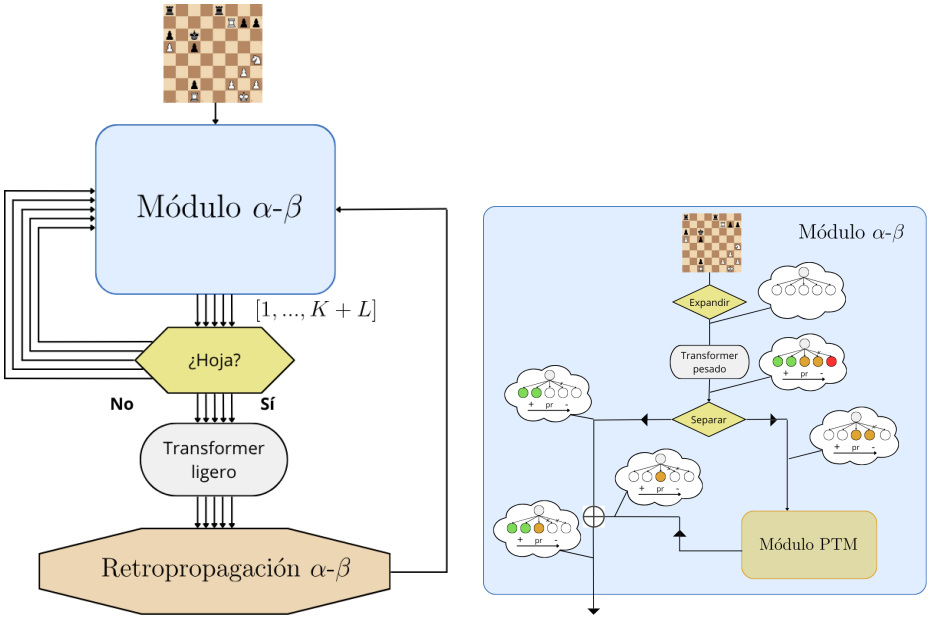
siendo este mecanismo conocido como *poda tardía de movimientos (PTM)*. La Figura 1 muestra, respectivamente, el flujo general y la lógica por nivel: en cada nodo no hoja, el transformer pesado evalúa la posición y se seleccionan los K mejores movimientos. A continuación, se hace una búsqueda para los siguientes W a profundidad 2 y, si alguno mejora, se incorpora al conjunto explorado.

3.2 Arquitectura con el mecanismo de reevaluación. DeliberaReev

Sea b el factor de ramificación y D la profundidad de la búsqueda alfa-beta. El número de nodos hoja crece a la razón de b^D. El transformer ligero es el modelo más rápido que actúa como evaluador de cada hoja, mientras que el pesado, más preciso pero más costoso, se usa en zonas puntuales del árbol. DeliberaReev reserva este último para posiciones más prometedores de acuerdo a las valoraciones del modelo rápido. Así, durante la subida en el árbol, cada nodo propaga el par (v, P), donde v es el valor y P el tablero del nodo hoja del que procede v, en lugar de solo devolver un escalar.

Dado que los niveles del árbol alternan entre nodos MAX y MIN según si juega el jugador o el rival, se aplican operaciones diferentes. En un nodo MAX (resp. MIN) se hereda el par del hijo de mayor (resp. menor) v, de modo que P es siempre la hoja que el modelo ligera considera el mejor en cada caso. La reevaluación se realiza al alcanzar cierta profundidad prefijada como hiperparámetro, momento en que el modelo pesado valora P, recalculando v y continuando con el flujo del algoritmo. Así, cada subárbol, con nodo raíz el nodo en que se reevalúa, usa una única vez el modelo pesado.

En este sentido, el mecanismo simula una búsqueda clásica, reproduciendo la operación que se haría con el modelo pesado como evaluador de hojas pero restringida a las posiciones que pasan el filtro del modelo ligero, evitando las b^D llamadas al modelo pesado. El flujo en general coincide con el de la versión base, con la única diferencia de que se arrastra el par (v, P). En la Figura 2 se observa lo explicado anteriormente.



(a) Flujo de DeliberaBasica que, dado un tablero, hace la búsqueda y evalúa las hojas con el transformer ligero.

(b) Flujo del algoritmo por nivel. Se lanza el transformer pesado (270M), obteniendo los K mejores movimientos. A continuación se exploran los siguientes W con PTM.

Figura 1: Arquitectura básica de la propuesta: flujo general (izquierda) y flujo por nivel (derecha).

4 Estudio experimental

En esta sección presentamos los resultados obtenidos sobre tres conjuntos de datos de prueba de naturaleza distinta. Mostramos tanto la configuración de los experimentos, las métricas obtenidas y la interpretación de las mismas.

4.1 Configuración de los experimentos

La implementación desarrollada y el marco experimental están disponibles en Github. [3].

Los experimentos se dividen en dos bloques: evaluación sobre conjuntos de puzzles y cálculo de Elo en partidas Blitz. Para el primero, contamos con tres datasets seleccionados con la supervisión de un Maestro FIDE [4]: dos conjuntos de puzzles de estrategia con dificultad media y alta (**SBP_MEDIUM**, 76

[3] https://github.com/ari-dasci/S-ThinkMore
[4] https://ratings.fide.com/profile/22245561

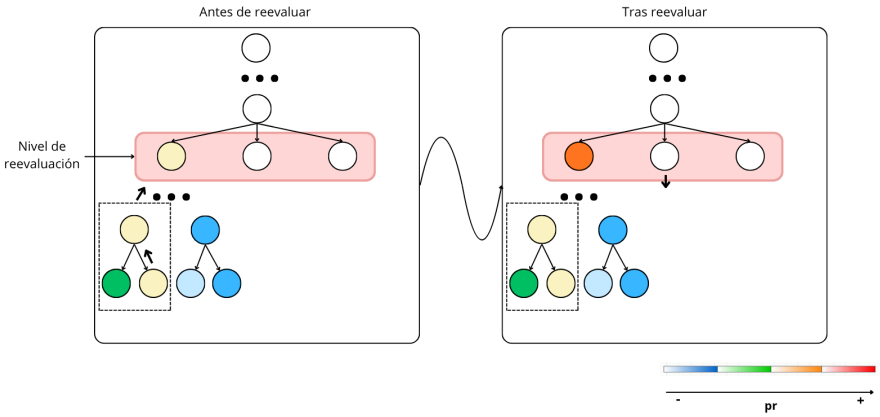


Figura 2: Representación gráfica del mecanismo de reevaluación.

posiciones, y **SBP_HARD**, 386 posiciones) [30, 27, 29] y un conjunto táctico de alta dificultad (**CBP_HARD**, 209 posiciones) [28]. Esta selección responde a la hipótesis de que los modelos puramente conexionistas rendirán mejor en escenarios estratégicos, donde el transformer internaliza principios heurísticos de contexto largo, mientras que su rendimiento caerá en táctica, que exige cálculo simbólico preciso y multi-paso.

Entre los hiperparámetros empleados se encuentra la profundidad de búsqueda (de 2 a 4), valor de poda K (de 4 a 6), tamaño de la ventana PTM, $W = 3$, y nivel de reevaluación fijado en la profundidad menos uno. Como transformer ligero se usan los modelos 9M, 136M y 270M para DeliberaBasica y 9M y 136M para DeliberaReev. El transformer pesado siempre será 270M.

La métrica principal es la precisión. En puzzles tácticos la solución es única, y dada la naturaleza de posibles movimientos subóptimos pero no malos, en los de estrategia se aceptará una solución alternativa siempre que la diferencia en probabilidad de victoria respecto a la referencia de Stockfish a profundidad 25 sea inferior a $\epsilon = 0.01$. Por otro lado, el Elo Blitz se calcula con 800 partidas contra el modelo 270M, aplicando la transformación

$$\Delta\text{Elo} = 400 \cdot \log_{10}\left(\frac{p}{1-p}\right)$$

[31], validado con el intervalo de Wilson al 95 % [32].

4.2 Resultados

Precisión en puzzles La Tabla 1 y la Figura 3 muestran los resultados a profundidad 2. La incorporación de búsqueda produce ganancias consistentes y sustanciales sobre los transformer base. Por ejemplo, DeliberaBasica(9M) pasa de 59.21 % a 80.26 % en SBP_MEDIUM, y DeliberaBasica(270M) mejora de 69.69 % a 80.83 % en SBP_HARD incorporando la búsqueda básica. Conforme aumenta

la profundidad, DeliberaReev muestra curvas más pronunciadas que DeliberaBasica, cuyas ganancias marginales se reducen al explotar la información relevante en profundidad baja.

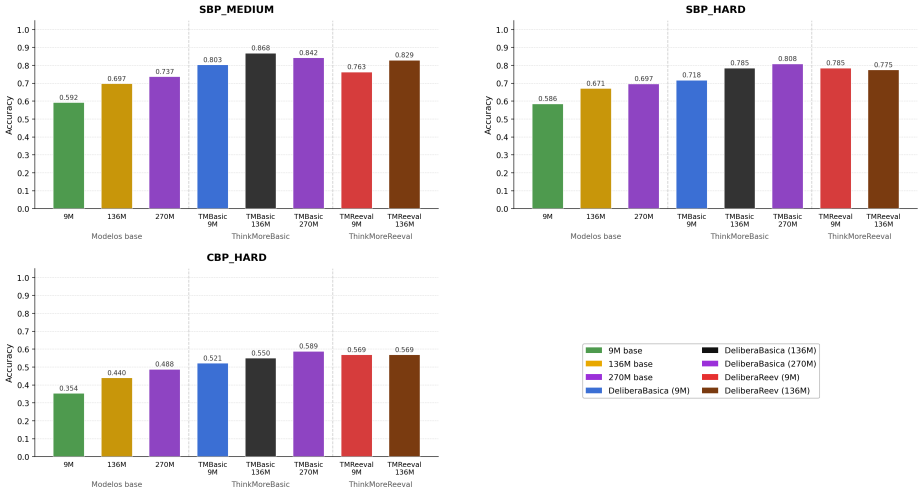


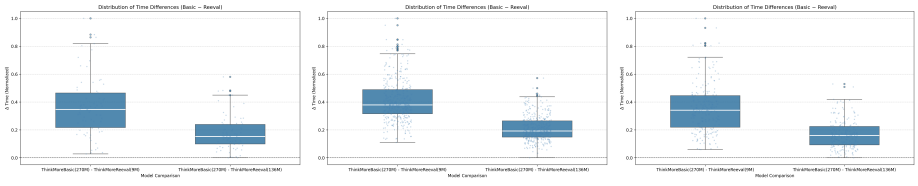
Figura 3: Representación gráfica de la precisión por dataset para todos los modelos (a profundidad 2 los que implementan búsqueda).

Modelo	SBP_MEDIUM	SBP_HARD	CBP_HARD
9M	0.5921	0.5855	0.354
136M	0.6973	0.671	0.4401
270M	0.7368	0.6969	0.488
DeliberaBasica(9M)	0.8026	0.7176	0.5215
DeliberaBasica(136M)	**0.8684**	0.785	0.5502
DeliberaBasica(270M)	0.8421	**0.8083**	**0.5885**
DeliberaReev(9M)	0.7632	0.785	0.5694
DeliberaReev(136M)	0.8289	0.7746	0.5694

Tabla 1: Precisión por dataset para todos los modelos con profundidad fija 2.

Eficiencia computacional La Figura 4 muestra la distribución de la diferencia de tiempo por puzzle entre la versión de búsqueda básica con 270M y las variantes con reevaluación. En todos los conjuntos de puzzles se observa que $\Delta t > 0$, confirmando que la reevaluación reduce consistentemente el tiempo de ejecución. Además, el ahorro es mayor con el transformer 9M, aunque la variante con

136M también mantiene un margen positivo estable. Esto completa la validación de la hipótesis, mostrando que el rendimiento es equivalente con menor coste computacional.



(a) SBP_MEDIUM (b) SBP_HARD (c) CBP_HARD

Figura 4: Distribución de la diferencia de tiempo de ejecución de la versión básica de búsqueda, DeliberaBasica(270M), respecto a las dos versiones que implementan el mecanismo de reevaluación.

Elo Blitz en Lichess La Tabla 2 recoge los resultados competitivos. La versión de DeliberaBasica(9M) obtiene un incremento de +21 puntos Elo sobre el modelo base 270M, aunque el intervalo de Wilson incluye el 50 % en tasa de victorias, por lo que la mejora no es estadísticamente significativa y la búsqueda simplemente actúa como reductor de la varianza. En contraste, DeliberaReev(9M) alcanza un Elo de 3075, con una ganancia de +180 puntos sobre el modelo base y +159 sobre la versión básica, con un intervalo de Wilson completamente desacoplado del umbral de paridad. Este resultado prueba que la reevaluación selectiva hace que el transformer sea un árbitro dinámico que dirige la búsqueda, logrando que un modelo de 9 millones de parámetros supere al de 270 millones usado de forma aislada.

Modelo	**Tasa victorias (%)**	**Elo estimado**	**Δ Elo**	**Wilson (%)**	**Wilson (Elo)**
270M (baseline)	—	2895	—	—	—
DeliberaBasica(9M)	53.0	2916	+21.0	[49.5, 56.4]	[−3, 45]
DeliberaReev(9M)	**73.8**	**3075**	**+180.0**	**[70.6, 76.7]**	**[152, 207]**

Tabla 2: Elo Blitz estimado en Lichess e intervalos de Wilson al 95% para las arquitecturas propuestas a profundidad 2.

5 Conclusiones y trabajo futuro

En este trabajo presentamos Delibera, un marco neurosimbólico que complementa el uso de transformers en una arquitectura estructurada de búsqueda

y reevaluación usando como campo de pruebas el ajedrez. La variante DeliberaBasica combina un transformer ligero como evaluador de nodos hoja con búsqueda minimax guiada por el modelo 270M, mientras que DeliberaReev hace una variación de este esquema con un mecanismo de reevaluación selectiva que reserva el modelo pesado para los puntos críticos del árbol. Los resultados obtenidos demuestran que ambas versiones superan de forma notable a los modelos base.

Este trabajo demuestra que la integración de modelos conexionistas con mecanismos de búsqueda simbólica constituye un paradigma superior al uso de transformers monolíticos de gran escala. La competencia no depende únicamente del número de parámetros, sino que también de una propiedad emergente de cómo estos se despliegan. La sinergia neurosimbólica propuesta combina el reconocimiento de patrones del transformer con la estabilidad deliberativa de la búsqueda estructurada, transformando la capacidad predictiva bruta en un proceso de decisión verificable que emula el razonamiento humano, donde se proponen candidatos de forma intuitiva y se someten a verificación sistemática.

Una contribución fundamental es la evidencia de que los mecanismos de reevaluación pueden sustituir eficazmente al escalado masivo de parámetros. La arquitectura jerárquica propuesta, donde un modelo de mayor capacidad actúa como supervisor táctico, no solo iguala, sino que supera a arquitecturas monolíticas mucho más pesadas, confirmando que el conocimiento de un gran modelo se aprovecha de forma más eficiente como evaluador de alto nivel que como motor de inferencia. En escenarios de alta complejidad, la robustez táctica depende más de la orquestación inteligente de modelos y la estructura lógica de la búsqueda que del número absoluto de parámetros.

Como trabajo futuro se plantean varias líneas. En primer lugar, en cuanto a la generalización del marco, se propone estudiar su aplicabilidad a dominios como la logística compleja, la demostración automática de teoremas o la planificación en entornos dinámicos, así como desarrollar un enfoque teórico en el que el modelo de alta capacidad se invoque únicamente cuando la incertidumbre colectiva de los modelos eficientes supere un umbral crítico.

En segundo lugar, respecto a la mejora del mecanismo de reevaluación, se plantea incorporar criterios de parada basados en la confianza del reevaluador, de modo que el sistema determine de forma autónoma cuándo una posición requiere evaluación adicional y cuándo la estimación inicial es suficiente.

Finalmente, en lo relativo al entrenamiento especializado, se propone implementar ciclos en los que el reevaluador aprenda específicamente de los fallos de los modelos en los nodos hoja, así como explorar el uso de otras arquitecturas en dichos nodos dentro de una misma búsqueda, permitiendo al sistema combinar distintas intuiciones antes de someterlas a la reevaluación.

Agradecimientos Esta investigación es parte del Proyecto TSI-100927-2023-1, financiado por el Plan de Recuperación, Transformación y Resiliencia de la Unión Europea–NextGenerationEU a través del Ministerio para la Transformación Digital y de la Función Pública, y está parcialmente financiada por el Ministerio de Ciencia, Innovación

y Universidades en el marco del proyecto PID2023-150070NB-I00, financiado por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

Declaración de conflicto de intereses Los autores declaran que no existe ningún conflicto de intereses en relación con la publicación de este artículo.

References

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) NIPS 2017, AnIPls, vol. 30, pp. 5998–6008. Curran Associates, Inc. (2017). https://doi.org/10.5555/3295222.3295349
2. Sapkota, R., Roumeliotis, K.I., Karkee, M.: AI agents vs. agentic AI: A conceptual taxonomy, applications and challenges. arXiv preprint arXiv:2505.10468 (2025), https://arxiv.org/abs/2505.10468, last accessed 2026/06/02
3. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In: International Conference on Learning Representations (ICLR 2021). OpenReview.net (2021), https://openreview.net/forum?id=YicbFdNTTy
4. Shi, C., Yang, K., Yang, J., Shen, C.: Transformers as Game Players: Provable In-context Game-playing Capabilities of Pre-trained Models. In: Advances in Neural Information Processing Systems (NeurIPS 2024), vol. 37, pp. 132001–132049. Curran Associates, Inc. (2024)
5. Ruoss, A., Delétang, G., Medapati, S., Grau-Moya, J., Wenliang, L.K., Catt, E., Reid, J., Lewis, C.A., Veness, J., Genewein, T.: Amortized planning with large-scale transformers: A case study on chess. In: Advances in Neural Information Processing Systems (NeurIPS 2024), vol. 37, pp. 65765–65790. Curran Associates, Inc. (2024)
6. Mira, J.M.: Symbols versus connections: 50 years of artificial intelligence. Neurocomputing 71(4-6), 671–680 (2008)
7. Edwards, D.J., Hart, T.P.: The alpha-beta heuristic. Tech. Rep., Massachusetts Institute of Technology, Cambridge, MA (1961)
8. Stockfish Developers: Stockfish: A Strong Open Source Chess Engine (2024), https://stockfishchess.org, version 17, last accessed 2026/06/02
9. Rosenblatt, F.: The Perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev. 65(6), 386–408 (1958). https://doi.org/10.1037/h0042519
10. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Nature 323(6088), 533–536 (1986). https://doi.org/10.1038/323533a0
11. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems (NIPS 2012), vol. 25, pp. 1097–1105. Curran Associates, Inc. (2012)
12. Achiam, J., Adler, S., Agarwal, S., et al.: GPT-4 technical report. arXiv preprint arXiv:2303.08774 (2023), https://arxiv.org/abs/2303.08774, last accessed 2026/06/02
13. Touvron, H., Lavril, T., Izacard, G., et al.: Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971 (2023), https://arxiv.org/abs/2302.13971, last accessed 2026/06/02

14. Guo, D., Yang, D., Zhang, H., et al.: DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning. arXiv preprint arXiv:2501.12948 (2025), https://arxiv.org/abs/2501.12948, last accessed 2026/06/02
15. OpenAI: Learning to Reason with LLMs (OpenAI o-series) (2024), https://openai.com/index/learning-to-reason/, last accessed 2026/02/12
16. Newell, A., Simon, H.A.: Computer Science as Empirical Inquiry: Symbols and Search. Commun. ACM 19(3), 113–126 (1976). https://doi.org/10.1145/1283920.1283930
17. Robinson, J.A.: A machine-oriented logic based on the resolution principle. J. ACM 12(1), 23–41 (1965)
18. Van Melle, W.: MYCIN: a knowledge-based consultation program for infectious disease diagnosis. Int. J. Man-Mach. Stud. 10(3), 313–322 (1978)
19. v. Neumann, J.: Zur Theorie der Gesellschaftsspiele. Math. Ann. 100(1), 295–320 (1928)
20. Trinh, T.H., et al.: Solving olympiad geometry without human demonstrations. Nature 625(7995), 476–482 (2024)
21. Chaudhuri, S., et al.: Neurosymbolic Programming. Found. Trends Program. Lang. 7(3), 158–243 (2021)
22. Turing, A.M.: Digital computers applied to games. In: Bowden, B.V. (ed.) Faster Than Thought, pp. 286–310. Pitman & Sons, London (1953)
23. Silver, D., Hubert, T., Schrittwieser, J., et al.: A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 362(6419), 1140–1144 (2018)
24. Coulom, R.: Efficient selectivity and backup operators in Monte-Carlo tree search. In: van den Herik, H.J., Ciancarini, P., Donkers, H.H.L.M. (eds.) Computers and Games (CG 2006). LNCS, vol. 4630, pp. 72–83. Springer, Berlin (2007)
25. Maharaj, S., Polson, N., Turk, A.: Chess AI: competing paradigms for machine intelligence. Entropy 24(4), Art. ID 550 (2022)
26. Goel, A.: Looking Back, Looking Ahead: Symbolic versus Connectionist AI. AI Mag. 42(4), 83–85 (2022). https://doi.org/10.1609/aaai.12026
27. Aagaard, J.: Grandmaster Preparation: Positional Play. Quality Chess, Glasgow (2013)
28. Aagaard, J.: Grandmaster Preparation: Calculation. Quality Chess, Glasgow (2013)
29. Aagaard, J.: Grandmaster Preparation: Strategic Play. Quality Chess, Glasgow (2013)
30. Bricard, E.: Strategic Chess Exercises: Find the Right Way to Outplay Your Opponent. New in Chess, Hilversum (2018)
31. Olesker-Taylor, S., Zanetti, L.: An Analysis of Elo Rating Systems via Markov Chains. In: Advances in Neural Information Processing Systems (NeurIPS 2024), vol. 37, pp. 138289–138323. Curran Associates, Inc. (2024)
32. Cook, J., Howard, J.J., Sirotin, Y., Tipton, J.L., Vemury, A.R.: Confidence Intervals for Error Rates in 1:1 Matching Tasks: Critical Statistical Analysis and Recommendations. IEEE Trans. Biometrics Behav. Identity Sci. 6(1), 110–123 (2024). https://doi.org/10.1109/TBIOM.2023.3331445