Imagen que contiene Logotipo

Descripción generada automáticamente

**Inteligencia Artificial**

****

**Investigación:**

**Agentes Deliberativos e Híbridos**

**Integrantes:**

**Armenta García Jonathan Humberto**

**Chairez Audelo Pedro**

**Dr. Zuriel Dathan Mora Felix**

**Ingeniería en Sistemas**

**Instituto Tecnológico de Culiacán**

**16/02/2025**

**Agentes Deliberativos**

**Definición:**

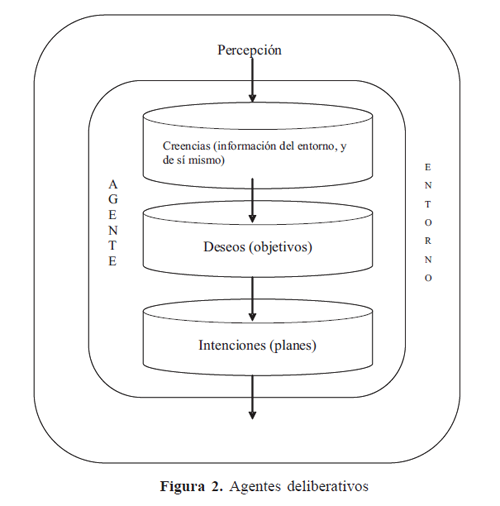
Los **agentes deliberativos** son un tipo de **agentes inteligentes** que toman decisiones basadas en un proceso de razonamiento lógico y planificación.

A **diferencia de los agentes reactivos**, que responden de manera inmediata a estímulos del entorno, los agentes deliberativos construyen un modelo del mundo, evalúan diferentes cursos de acción y eligen el mejor según sus objetivos.

**Arquitectura:**

Una de las arquitecturas más utilizadas en agentes deliberativos es **BDI**, que se basa en tres conceptos fundamentales:

* **Beliefs (Creencias)**: Lo que el agente sabe sobre el mundo y sobre sí mismo.
* **Desires (Deseos)**: Los objetivos que el agente quiere alcanzar.
* **Intentions (Intenciones)**: Los planes de acción que el agente decide ejecutar para cumplir sus deseos.



Pasos de la arquitectura BDI:

1. El agente **percibe** el entorno y actualiza sus creencias.
2. Evalúa qué deseos puede perseguir en función de sus creencias.
3. Selecciona una intención y elabora un plan de acción.
4. Ejecuta el plan y vuelve al ciclo de percepción.

**Características:**

* Mantienen una representación del entorno en el que operan.
* Evalúan posibles acciones y seleccionan la más adecuada.
* No actúan de forma impulsiva, sino que buscan cumplir objetivos a largo plazo.
* Pueden modificar su plan de acción en función de cambios en el entorno.
* Aplican reglas, heurísticas o incluso técnicas de IA como el razonamiento probabilístico para la toma de decisiones.

**Problemas:**

* Requieren gran capacidad de procesamiento debido a la necesidad de evaluar múltiples opciones antes de actuar.
* Su razonamiento deliberativo puede ser más lento que los agentes reactivos, lo que dificulta su uso en entornos donde la rapidez es crítica.
* Desarrollar un modelo del mundo preciso y eficiente no es trivial.
* En entornos dinámicos, los cambios pueden invalidar los planes previstos, requiriendo adaptación continua.

**Ejemplos:**

* **Vehículos autónomos** con sistemas como los de Tesla que utilizan modelos del mundo, predicen el comportamiento de otros vehículos y planifican su trayectoria.
* **Asistentes virtuales avanzados** como ChatGPT o Google Bard que pueden razonar múltiples preguntas y planificar respuestas complejas.
* **Robots de exploración espacial**: El rover Perseverance de la NASA toma decisiones sobre qué rocas analizar y cómo moverse por terrenos difíciles.
* **Agentes en videojuegos de estrategia**: En juegos como StarCraft II, los agentes de IA planifican estrategias a largo plazo en lugar de solo reaccionar a eventos inmediatos.
* **Sistemas de recomendación avanzada** en plataformas como Netflix o Amazon empleando modelos para predecir qué contenido será más adecuado para cada usuario.

**Agentes Deliberativos**

**Definición:**

Los agentes híbridos son sistemas de inteligencia artificial que combinan diferentes paradigmas o enfoques de IA, como la lógica simbólica y las redes neuronales, para resolver problemas complejos de manera más eficiente. Su principal fortaleza radica en su capacidad para aprovechar las ventajas de cada enfoque, superando las limitaciones de utilizarlos de manera aislada.

**Arquitectura:**

La arquitectura de los agentes híbridos se diseña para integrar múltiples componentes que colaboran entre sí. Algunos elementos clave son:

**Componentes Principales:**

* **Motor simbólico:** Utiliza reglas lógicas para razonar y tomar decisiones basadas en conocimiento explícito.
* **Red neuronal:** Aprende patrones a partir de datos y generaliza soluciones para problemas complejos.
* **Mecanismo de integración:** Coordina la interacción entre los componentes para lograr objetivos comunes.

**Modelos Comunes:**

* **División por tareas:** Cada componente se especializa en una tarea específica.
* **Comunicación centralizada:** Existe un controlador principal que gestiona el flujo de información.
* **Arquitectura jerárquica:** Los componentes están organizados en niveles para facilitar la colaboración.

**Características:**

* **Flexibilidad:** Capacidad para abordar diferentes tipos de problemas.
* **Adaptabilidad:** Mejora continua gracias al aprendizaje automático.
* **Robustez:** Combina la precisión de los sistemas simbólicos con la capacidad de generalización de las redes neuronales.
* **Escalabilidad:** Puede ajustarse a problemas de mayor complejidad mediante la adición de componentes.
* **Interoperabilidad:** Fácil integración con otras tecnologías.

**Problemas:**

**Complejidad en el Diseño**

* Diseñar y coordinar la interacción entre los componentes puede ser desafiante.

**Costos Computacionales**

* Requiere recursos significativos para entrenar y operar los diferentes subsistemas.

**Conflictos de Decisión**

* Integrar decisiones tomadas por paradigmas diferentes puede generar inconsistencias.

**Dificultad en la Explicabilidad**

* Los resultados de un sistema híbrido pueden ser difíciles de interpretar debido a la combinación de enfoques.

**Ejemplos de Uso:**

**Asistentes Virtuales**

* Combinan procesamiento de lenguaje natural simbólico y aprendizaje profundo para mejorar la interacción con usuarios.

**Diagnóstico Médico**

* Integran bases de conocimiento médico (simbólico) con sistemas de detección de patrones (redes neuronales) para diagnosticar enfermedades.

**Sistemas de Conducción Autónoma**

* Utilizan modelos simbólicos para reglas de tráfico y redes neuronales para la percepción y toma de decisiones en tiempo real.

**Videojuegos**

* Crean agentes no jugables (NPCs) que combinan lógica para estrategias y redes neuronales para comportamientos adaptativos.