



ROBOTIC WARFARE

la robótica en la guerra

¿Qué temas trataremos?



Las dificultades que pueden producirse al automatizar la guerra.



¿Por qué ocurren estos problemas?



Posibles soluciones.

ESCENARIOS EN LA GUERRA

Combatientes

No combatientes



- ▶ **Adscripción del estado mental.**
Requerido mayoritariamente en escenarios con mayoría no combatiente.
- ▶ **Isotropía.**
La relevancia potencial de cualquier cosa a cualquier cosa.





PHALANX

- ▶ Automatización de la selección de objetivos
(no distingue entre amigo y enemigo).
- ▶ CIWS (close in weapon system).
- ▶ Creado y testado en los 70s.
- ▶ Usado por los marines hoy en día.
- ▶ Supresión manual.
- ▶ Mayoritariamente usado en escenarios con combatientes.

MISIÓN: OPERACIÓN DE CONTRAINSURGENCIA

- ▶ Las fuerzas terrestres reciben un soplo erróneo de unos insurgentes tomando asilo en una residencia de civiles (La unidad no lo sabe).
- ▶ Tres niños y sus padres están presentes en la residencia.
- ▶ Dos de los niños están jugando con una pelota mientras llevan en mano un kirpan.
- ▶ Uno de los niños pateo la pelota hacia la puerta, y ambos niños corren hacia ella.
- ▶ Los militares entran en la casa y ven dos niños corriendo hacia ellos, y una madre gritándoles y persiguiéndoles.
- ▶ La unidad no sabe que está diciendo ya que no entienden su idioma.

INTERPRETACIONES

- ▶ Dos niños jugando.
 - ▶ Una madre asustada por sus hijos.
-

- ▶ Dos posibles objetivos armados acercándose.
- ▶ Un tercer objetivo causando ruido amenazante.



La atribución de estados mentales

La atribución de estados mentales es la asignación de creencias, sentimientos, esperanzas, miedos, intenciones, etc..

Hay dimensiones descriptivas y dimensiones normativas en el estudio de nuestras habilidades cotidianas para:

- ▶ Atribuir creencias, esperanzas, miedos,...
- ▶ Predecir lo que un agente hizo o hará.

La atribución de la intención

Escenarios en los que **no es necesario** determinar lo que el objetivo **piensa o siente** en este caso → **Escenario combatiente**.
(Buque de guerra con sistema antimisil)

Escenarios en los que **no podemos asumir** que todos son **combatientes** → **Escenario con posibilidad de civiles**.
(El ejemplo de los niños y el kirpan)

¿Qué problemas tenemos que solucionar para poder diseñar sistemas que puedan atribuir estados mentales confiables en este ultimo contexto?

Isotropía

La isotropía hace referencia a la potencial relación de las cosas con las cosas.

Ejemplo:



Theory Theory VS Simulación Theory

Simulación Theory(ST): Consiste en el entendimiento de otra persona mediante la recreación o copia de los estados mentales de esta persona en uno mismo.

Theory Theory(TT): Requiere que se tengan representaciones generalizadas(como reglas, estructuras,...) que corresponden a las supuestas conexiones entre los estados mentales y las acciones.

Theory Theory VS Simulación Theory

Si corremos una simulación basada en las acciones propias, estados mentales y procesos (ST), se nos quedaría el modelo insatisfascible por dos razones:

1. Se necesitan construir los mecanismos para llegar a los diferentes posibles estados mentales antes de poder simular otros.
2. Aunque se consigan construir los mecanismos, por culpa de la isotropía diferentes robots y/o personas llegarán a diferentes estados mentales con la misma información de entrada.

Emociones

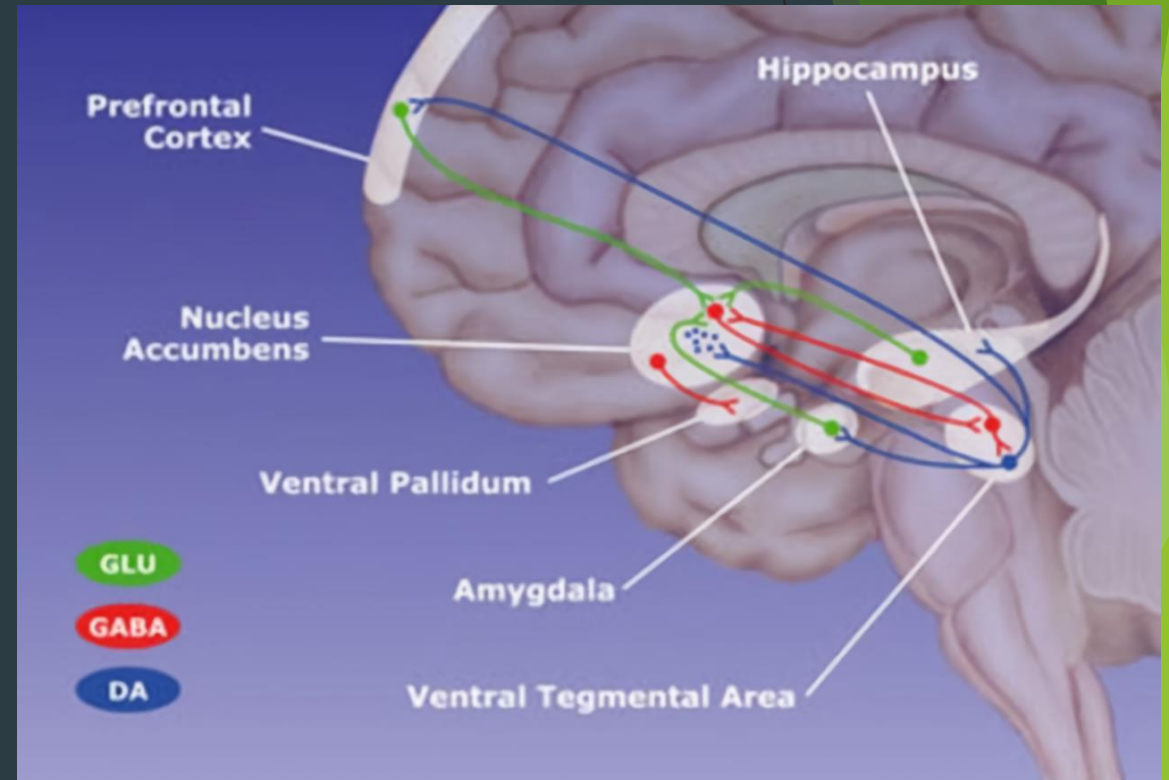
Gobernador Ético: incluye un rol limitado para las emociones.

Una idea de implementación podría ser:

- ▶ Introduce una función de un **rol** equivalente a la **'culpa'**, el cual se va modificando con las acciones y sus consecuencias.
- ▶ El valor de **'culpa'** se puede reiniciar con **asesoramiento externo**.
- ▶ El valor **censura o veta acciones** en escenarios parecidos a los ya ocurridos.

Modelo de decisiones

Wagar y Thagar implementan funciones del corteza prefrontal ventromedial (VMPFC), hipocampo, amígdala, núcleo accumbens, tegmento ventral para tomar decisiones.



¿VMPFC dañada?

Efectos:

- ▶ Dificultad en tomar decisiones efectivas.
- ▶ Filtro establecidos con anterioridad dañados o eliminados.
- ▶ Imposibilidad de crear nuevos filtros.
- ▶ Falta de empatía.
- ▶ Dificultad para navegar en entornos sociales.

Domar la isotropía

Consideramos un sistema

Objetivo: Hacer inferencias relevantes y evitar las irrelevantes.

Suponemos que:

- ▶ Está equipado con un foco de atención que almacena una proposición cada vez.
- ▶ Un subsistema coge una sola proposición P .
- ▶ Coge un set activo de proposiciones S como entrada.
- ▶ Devuelve un valor escalar $E \in (0,1)$ como salida.
- ▶ Dispone de un conocimiento previo aceptable K .

Domar la isotropía

Capacidades del sistema

- ▶ Las proposiciones pueden ser generadas internamente o por percepciones.
 - ↳ Las proposiciones pueden ser llamadas desde memoria o generadas por percepciones recogidas de un sensor de datos.
- ▶ Es capaz de adoptar creencias, deseos, intenciones,...
- ▶ Es capaz de planear y usar 'mindreading'.
- ▶ Tiene un mecanismo de evaluación que genera variables urgentes para cada motivación, deseo y objetivo.

Domar la isotropía

Funcionamiento:

- ▶ En cada ciclo cognitivo se genera E desde el foco actual P.
- ▶ La representación de la situación S y los valores de urgencia son generados por motivaciones, deseos y objetivos.
- ▶ La próxima proposición en tener el foco de atención será el resultado de uno de los escalares(E o una de las fuentes de los valores de urgencia).
- ▶ Definimos el espacio del problema en el cual este tipo de motor de inferencias necesita operar-> $K+S$

¿Solución al problema de la isotropía?

Si tenemos que:

- ▶ P es el foco de atención actual
- ▶ El subsistema emocional genera un valor escalar lo suficiente mente alto E para P , S , y sus consecuencias inferenciales inmediatas.

Entonces:

- ▶ El proceso de gestión de la atención o se quedaría en P o se movería a P^* .

Siendo P^* :

- 1.-Semánticamente cercana a P .
- 2.-Una consecuencia emocional relevante de P y S .

P sería entonces un imán de atención.

Conclusión

