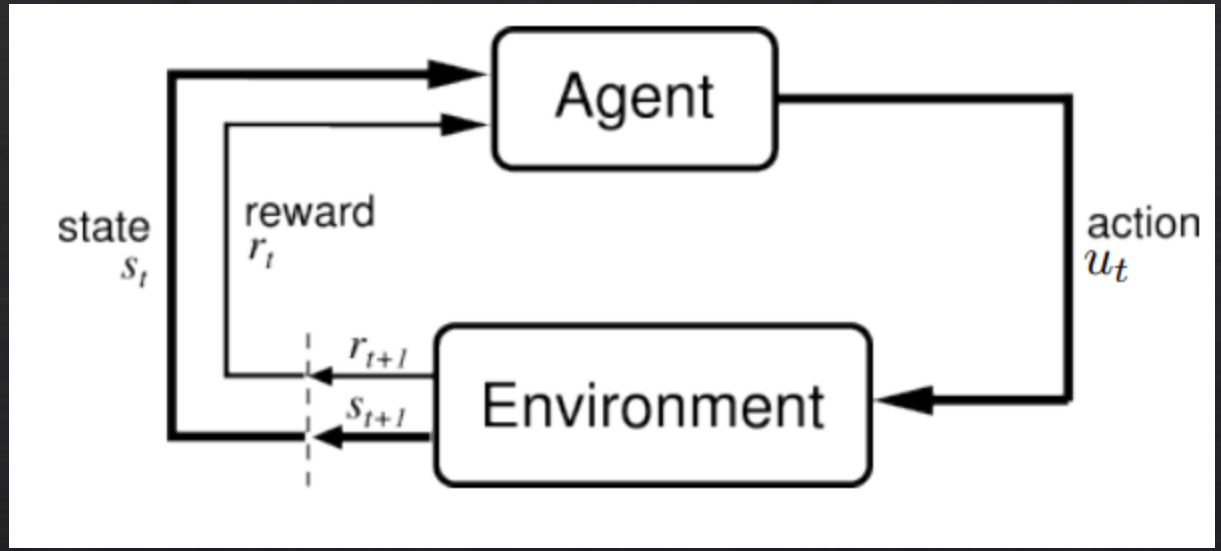
# Clase 6.3

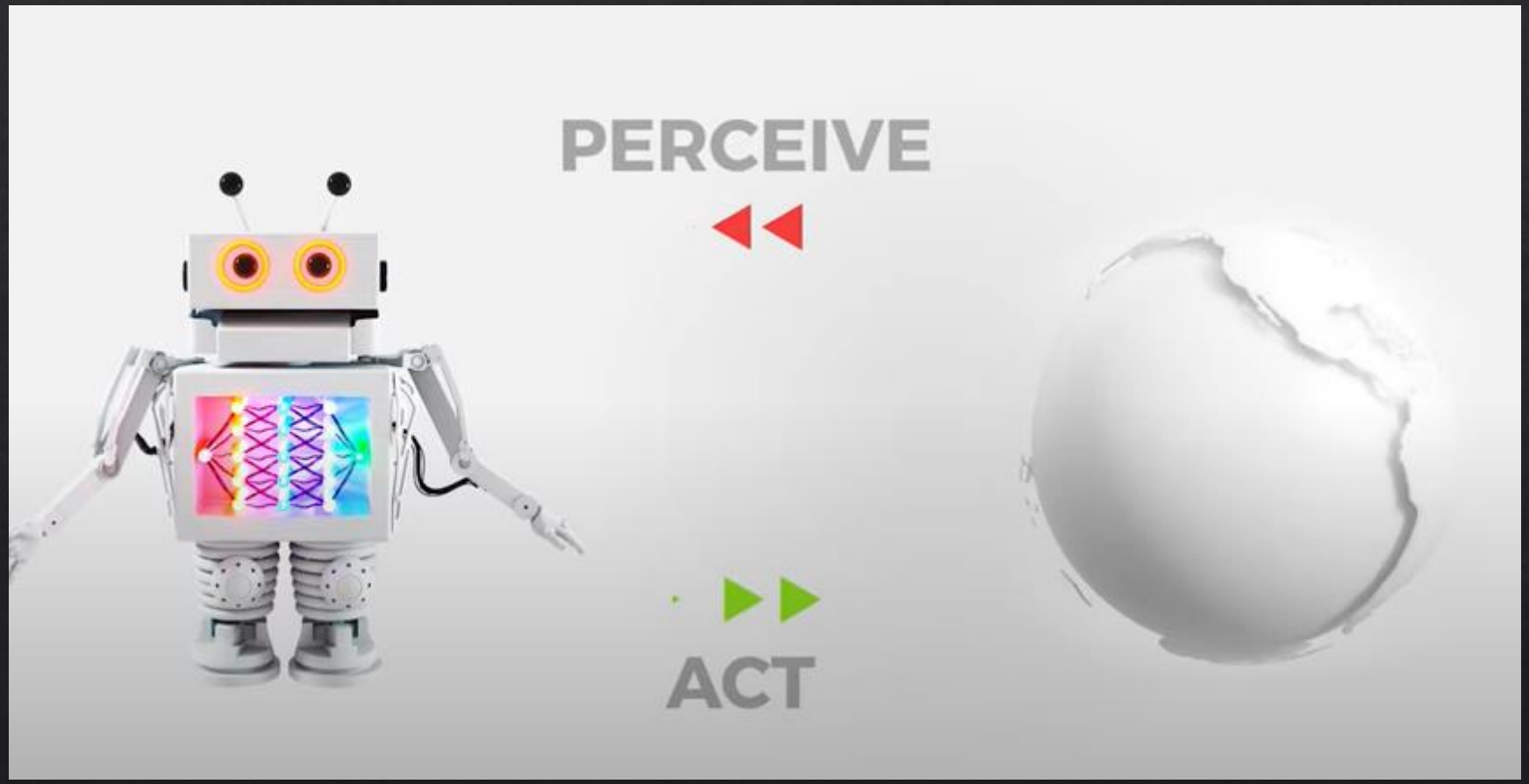
Aplicaciones de RL

Jorge Vásquez

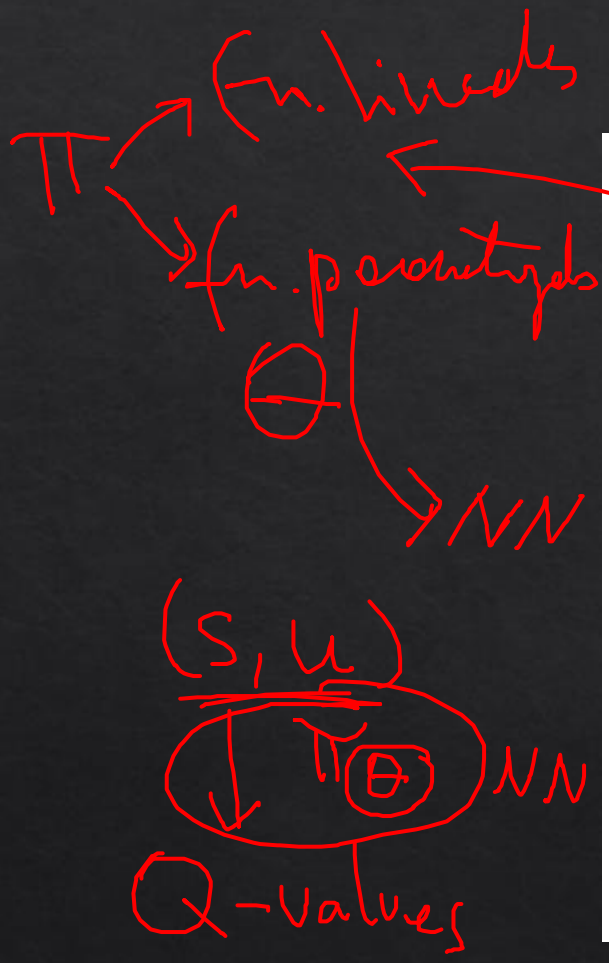
# RL

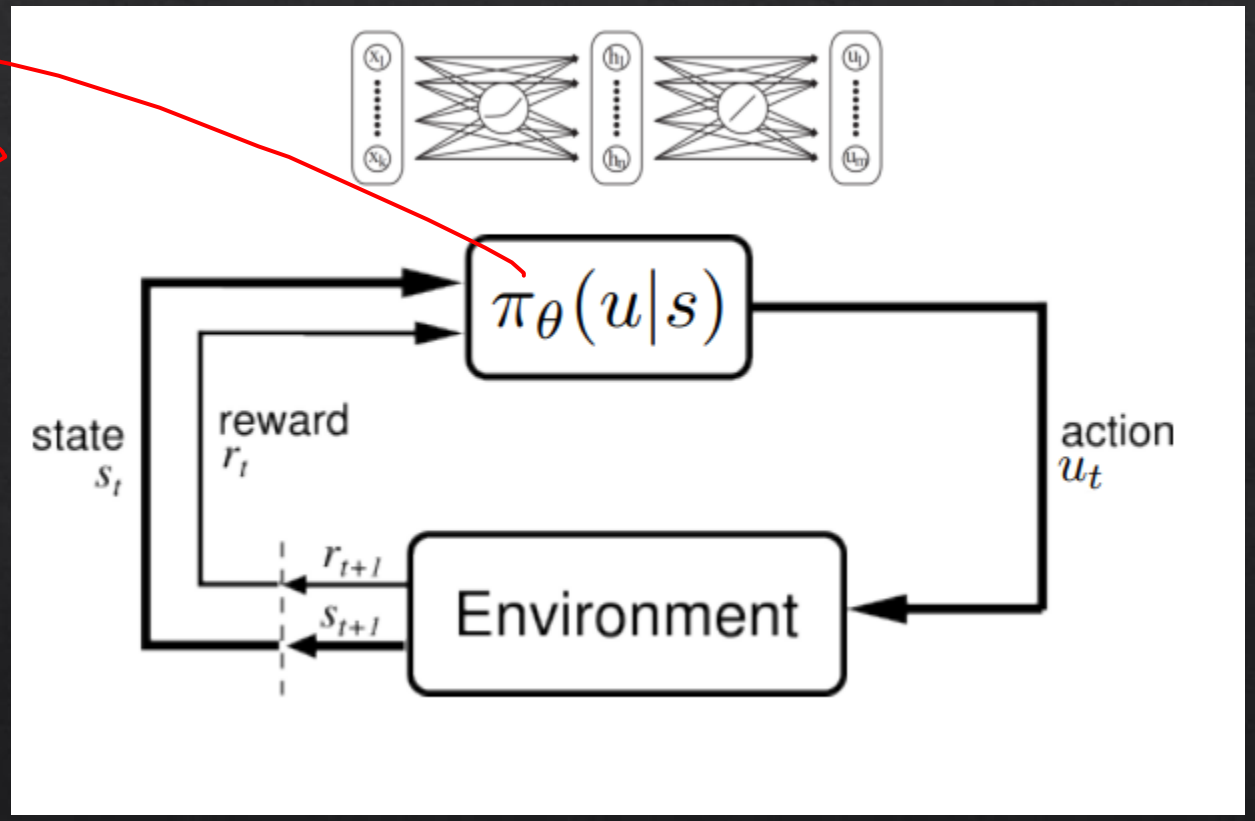


# ¿Cómo puede ayudar el Aprendizaje Reforzado?

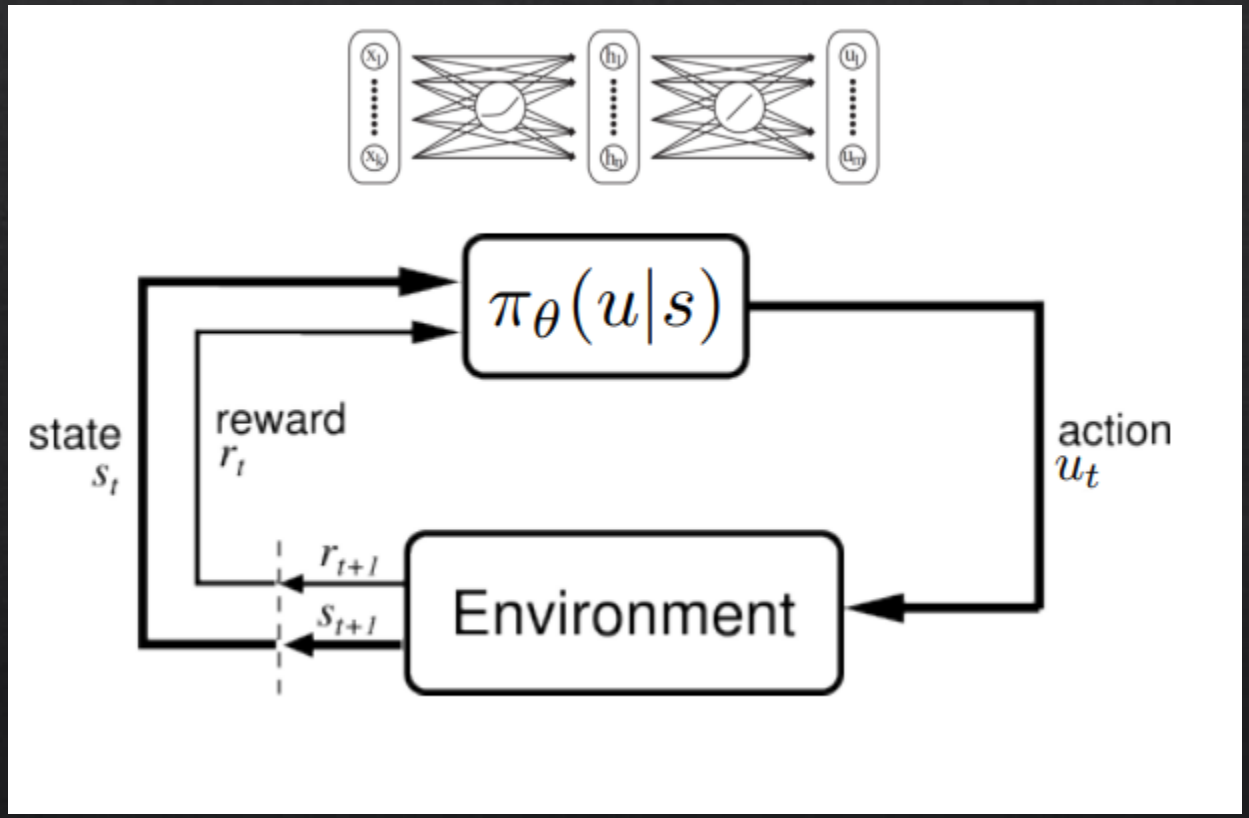


# DRL para Optimizar la Política





# DRL para Optimizar la Política




$$\max_{\theta} \quad \mathrm{E}[\sum_{t=0}^{H} R(s_t)|\pi_\theta]$$

# DRL para Optimizar la Política, ejemplos de apps

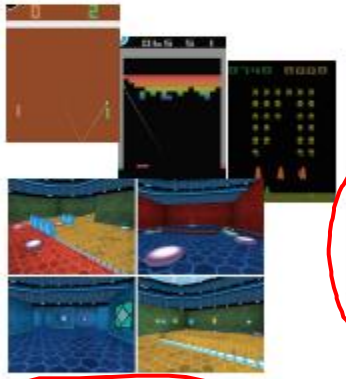
Kohl and Stone, 2004

Ng et al, 2004

Andrew Ng

Tedrake et al, 2005

Kober and Peters, 2009



Mnih et al, 2015 (A3C)

Silver et al, 2014 (DPG)
Lillicrap et al, 2015 (DDPG)

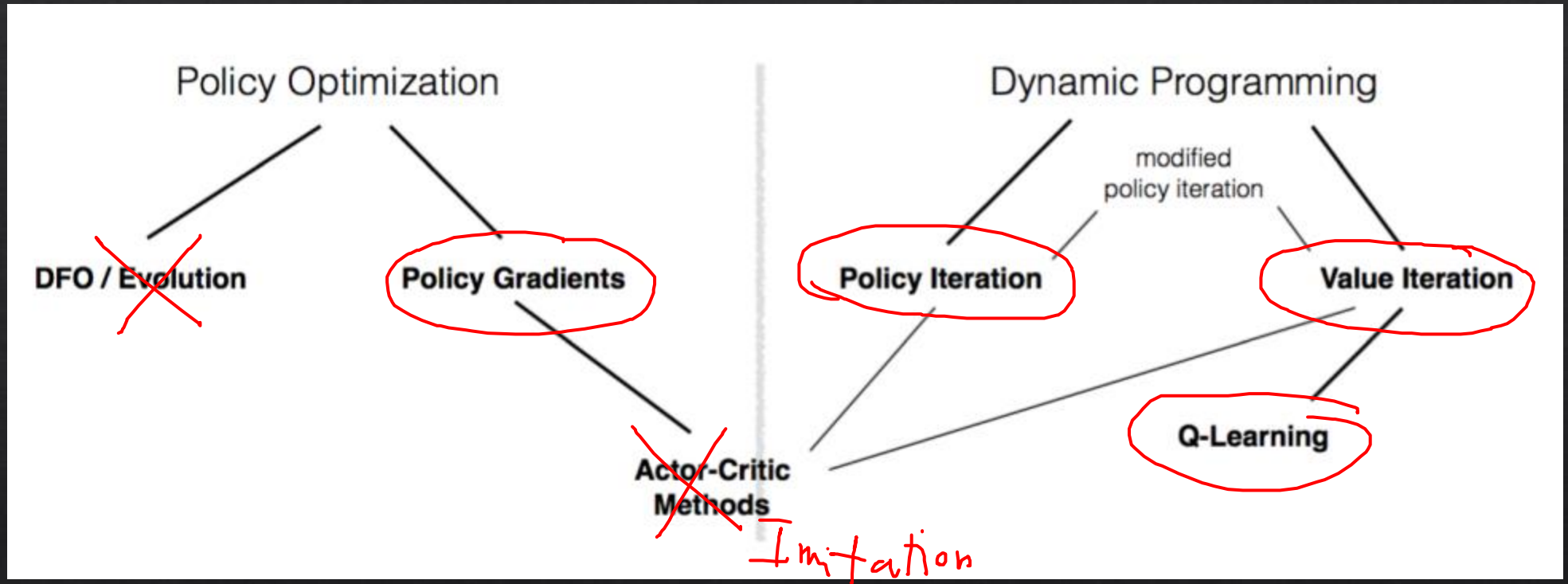

Schulman et al, 2016 (TRPO + GAE)

Abbel

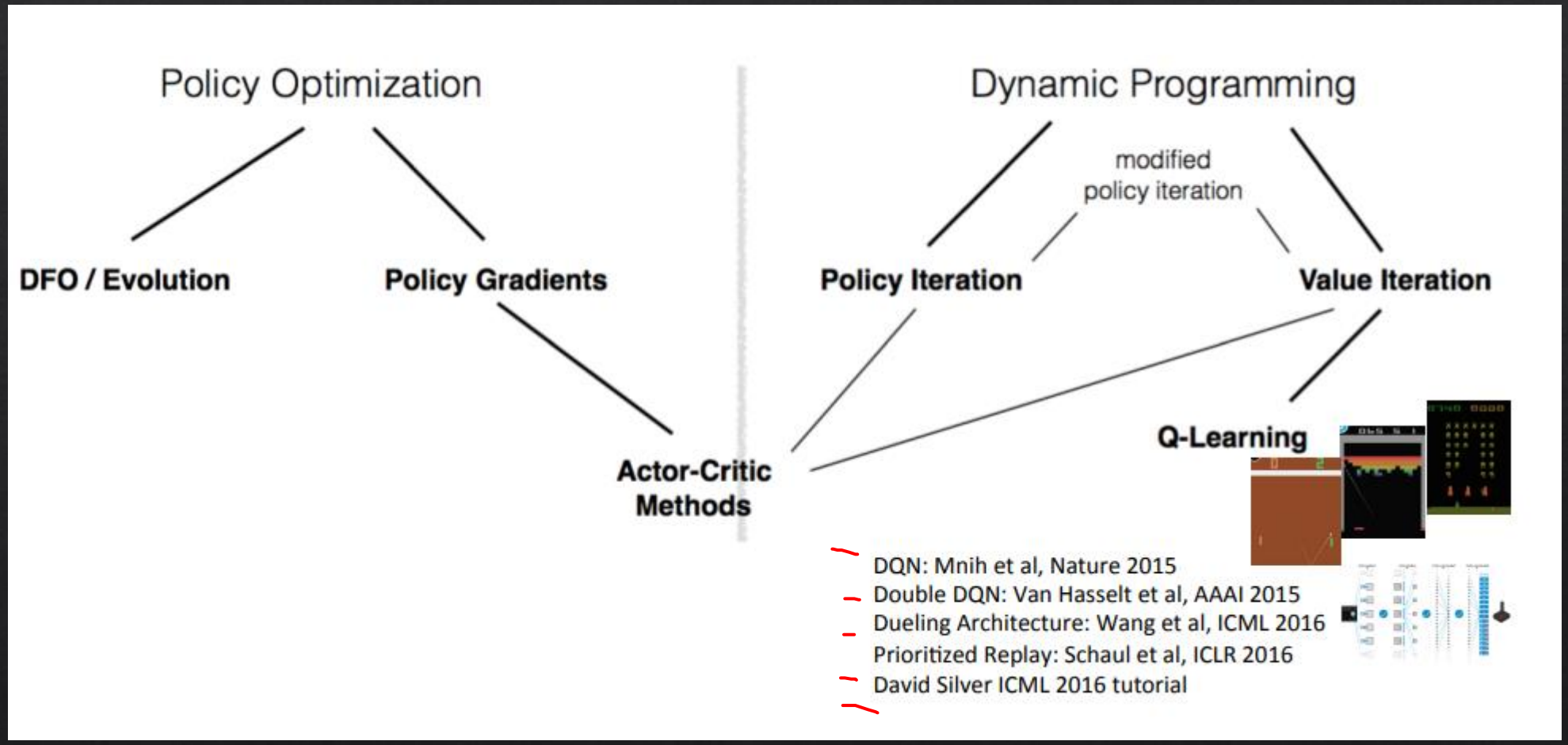
Levine*, Finn*, et al, 2016 (GPS)


Silver*, Huang*, et al, 2016 (AlphaGo**)

# RL para Optimizar la Política

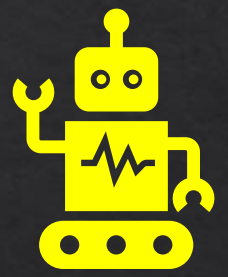


# RL para Optimizar la Política



# Diferencias entre investigación y mundo real

Academia



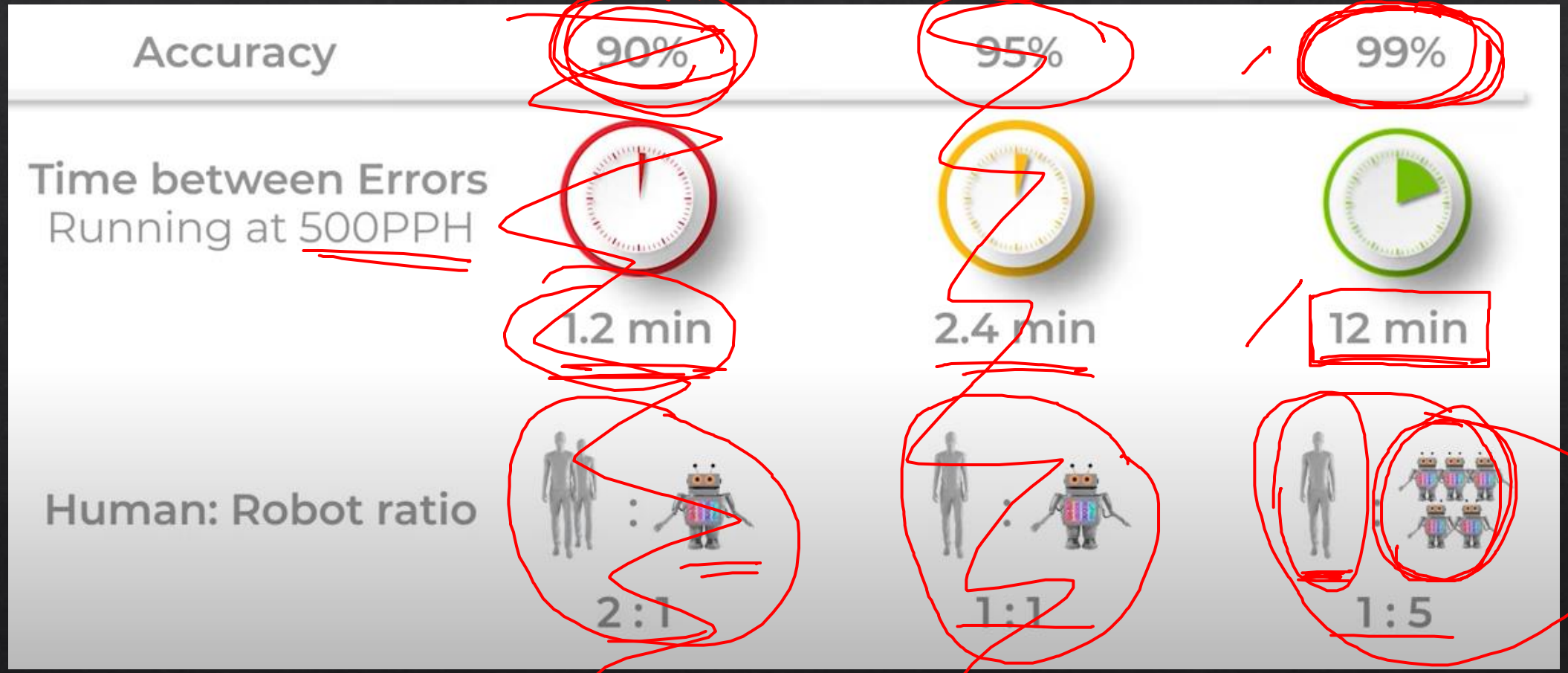
- De la nada a algo 0 → 1
- Al 70% de confianzas, nos movemos a otro problema

VS

Industria

- 90% no es suficiente

# Diferencias entre investigación y mundo real



# Problemas en el mundo real

1. No puedes ignorar la cantidad de variables en entornos reales
2. No puedes ignorar que el entorno es dinámico
3. No puedes ignorar lo que no sabes

Amazon

Kiva Systems

- Ejército de Robots

- 0.7B

- 70% Comprobar
  30% Pick & Place

→ Prime, 2012

# Problemas en el mundo real

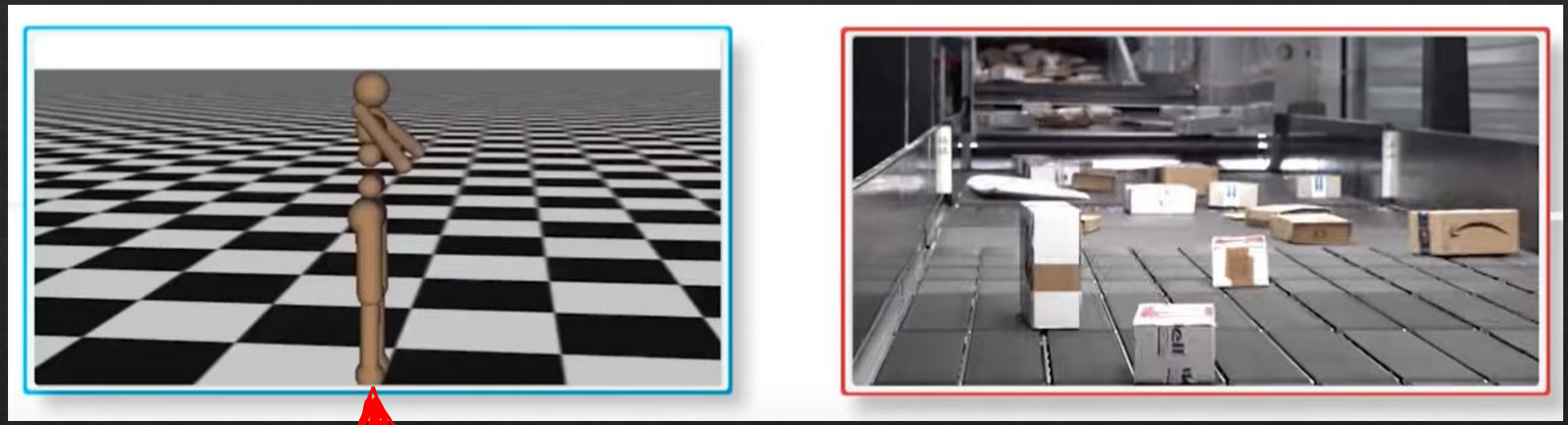
1. No puedes ignorar la gran cantidad de variables en entornos reales


- 1000 categorías
- 1000 ejemplos por categorías

- Millones de SKU
- Transparencia
- Cambios no conocidos

# Problemas en el mundo real

1. No puedes ignorar que el entorno es dinámico



~90%

- Simulación
- Se entrena previamente

S2RW

SimToReal

- Adaptarse en el momento

fallas
sensores
clima

# Problemas en el mundo real

1. No puedes ignorar lo que no sabes

# Problemas en el mundo real

1. Cuando empezaran los autos autonomos_?

# ¿Cuándo empezamos?

CoVariant → Abbel

Knapp


- Manipulación
→ Grasping

https://www.dcvelocity.com/media/videos/play/2285-knapp-pick-it-easy-robot-powered-by-covariant

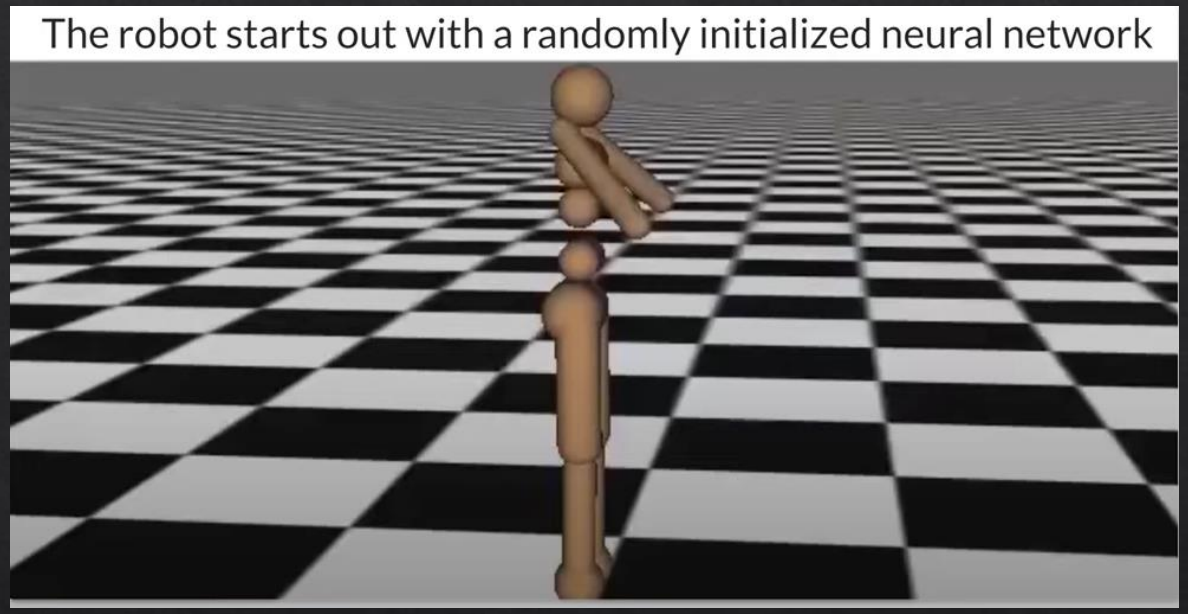
# 1. Videos Juegos, DeepAI

# 2. Robótica para la Inspección

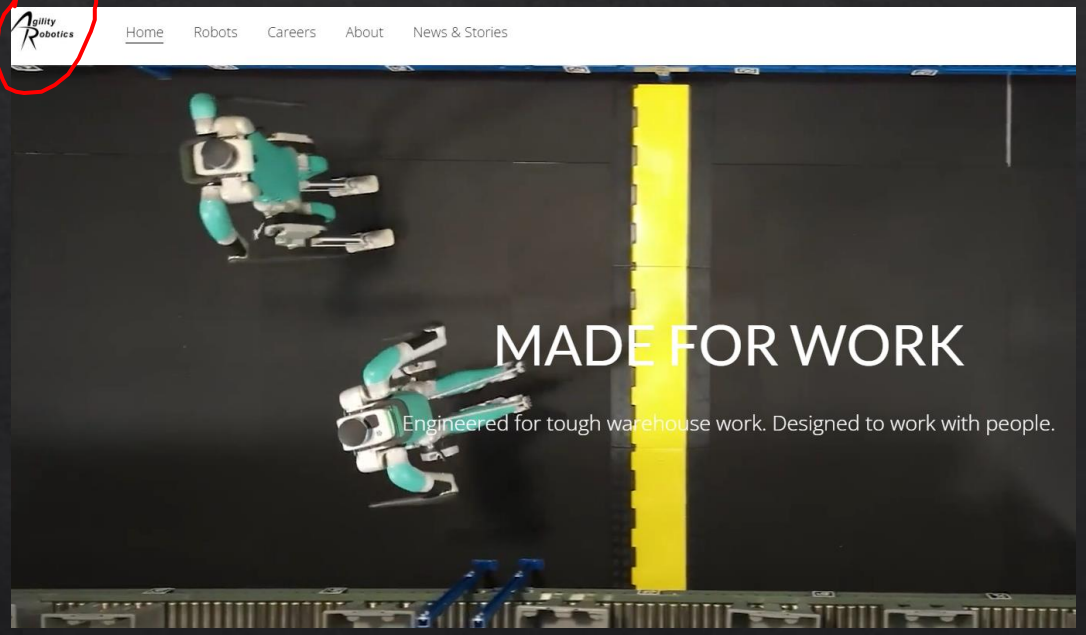
RL para Locomoción

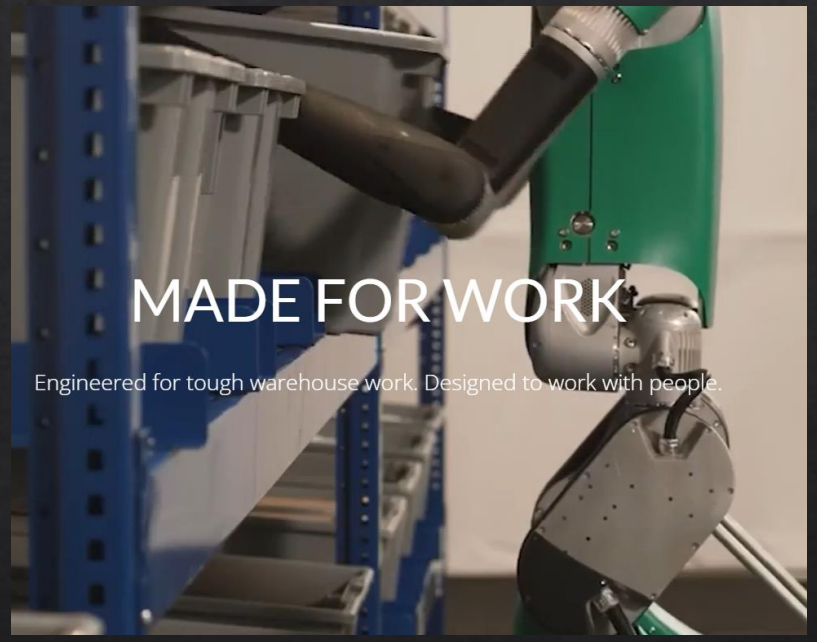
# 2. Aprender a moverse

Locomotion



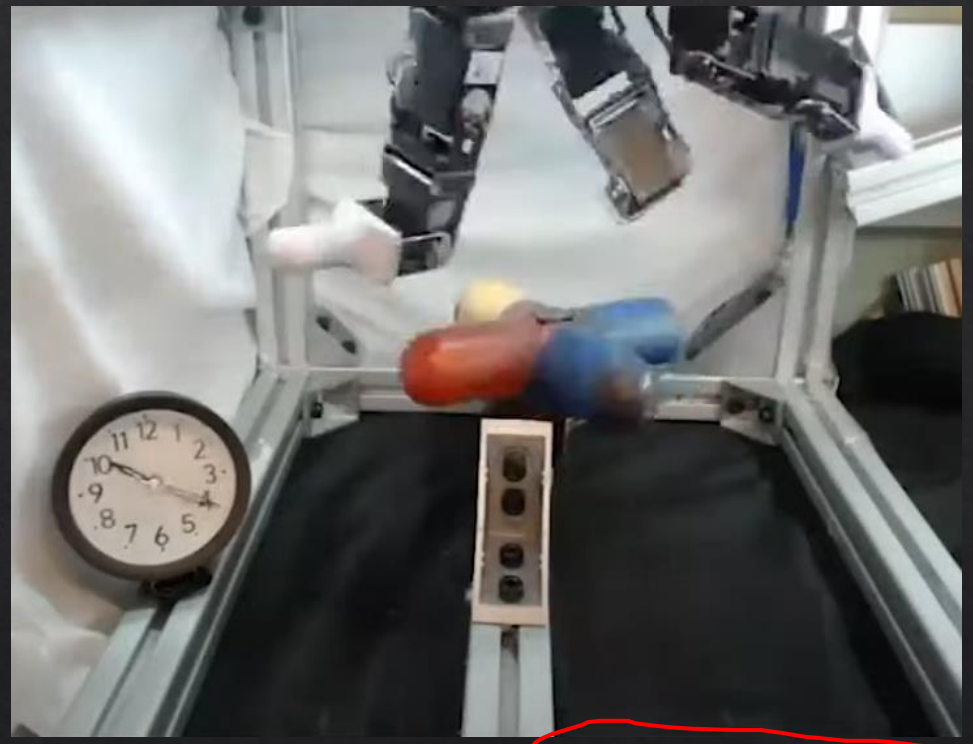
# 3. Humanoides

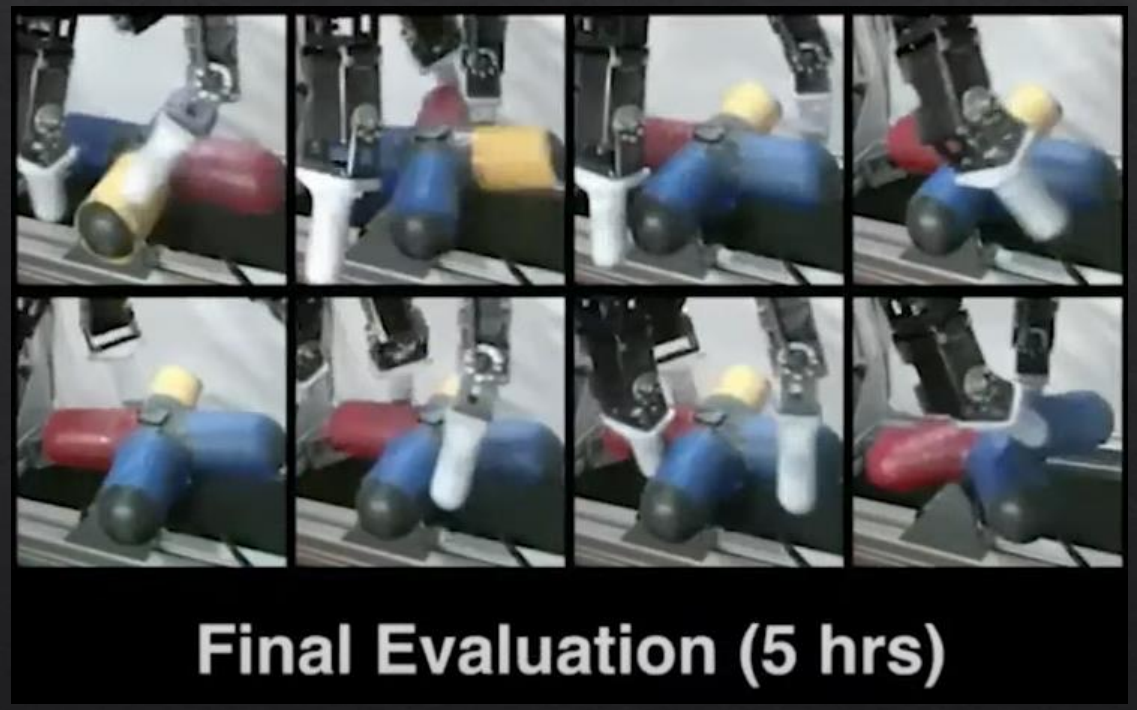


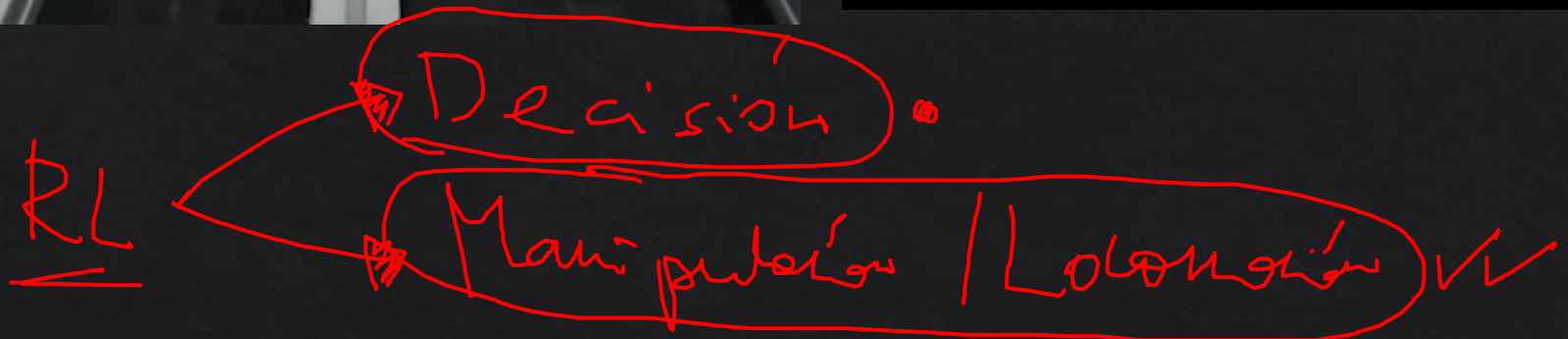




- Pick & Place

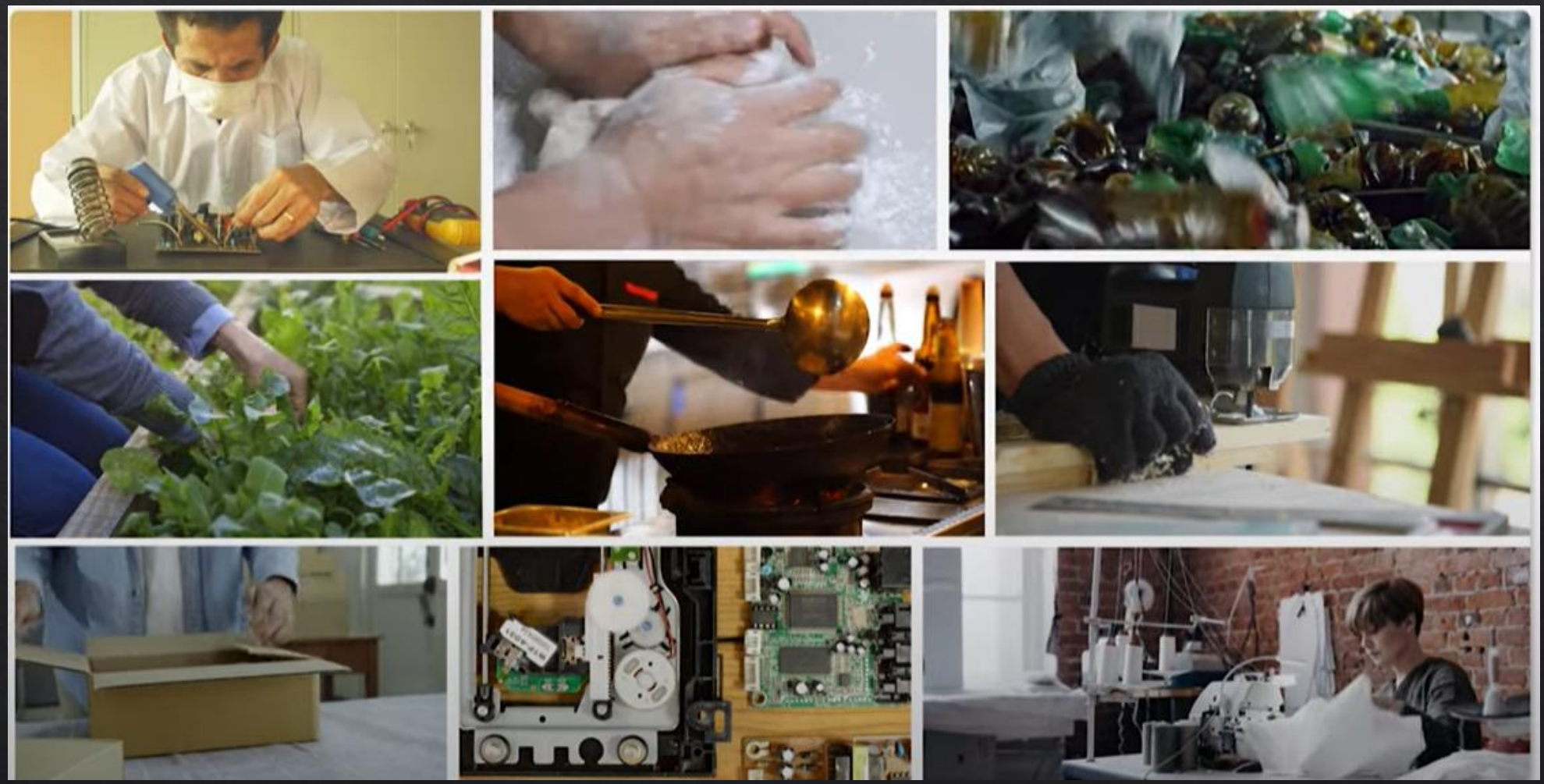
# 4. Manipulación Robótica



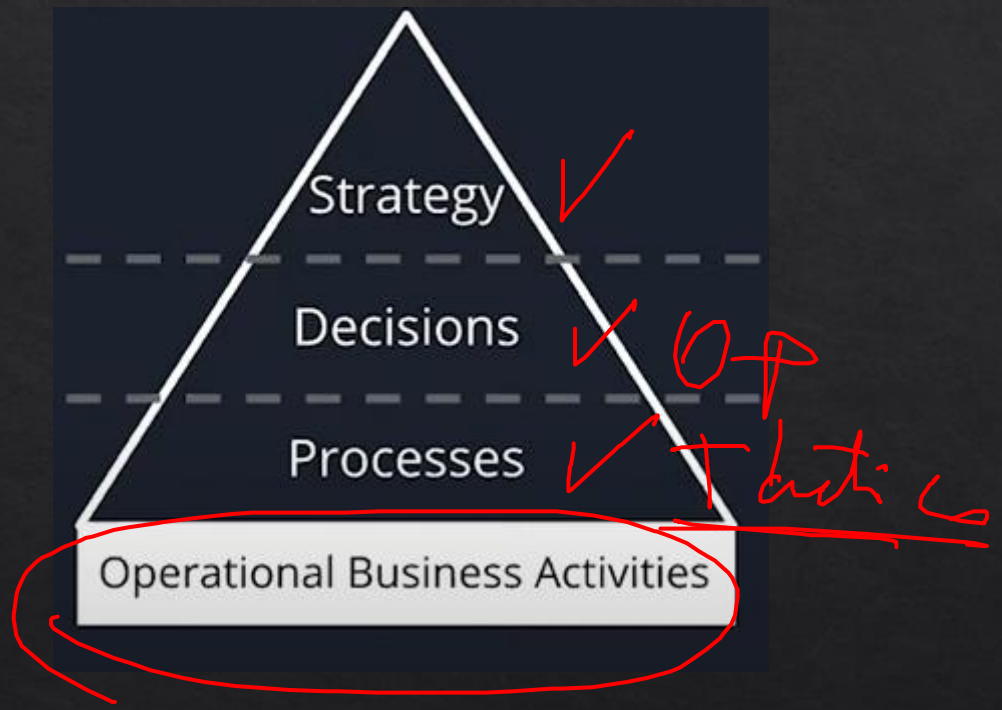


# Otros

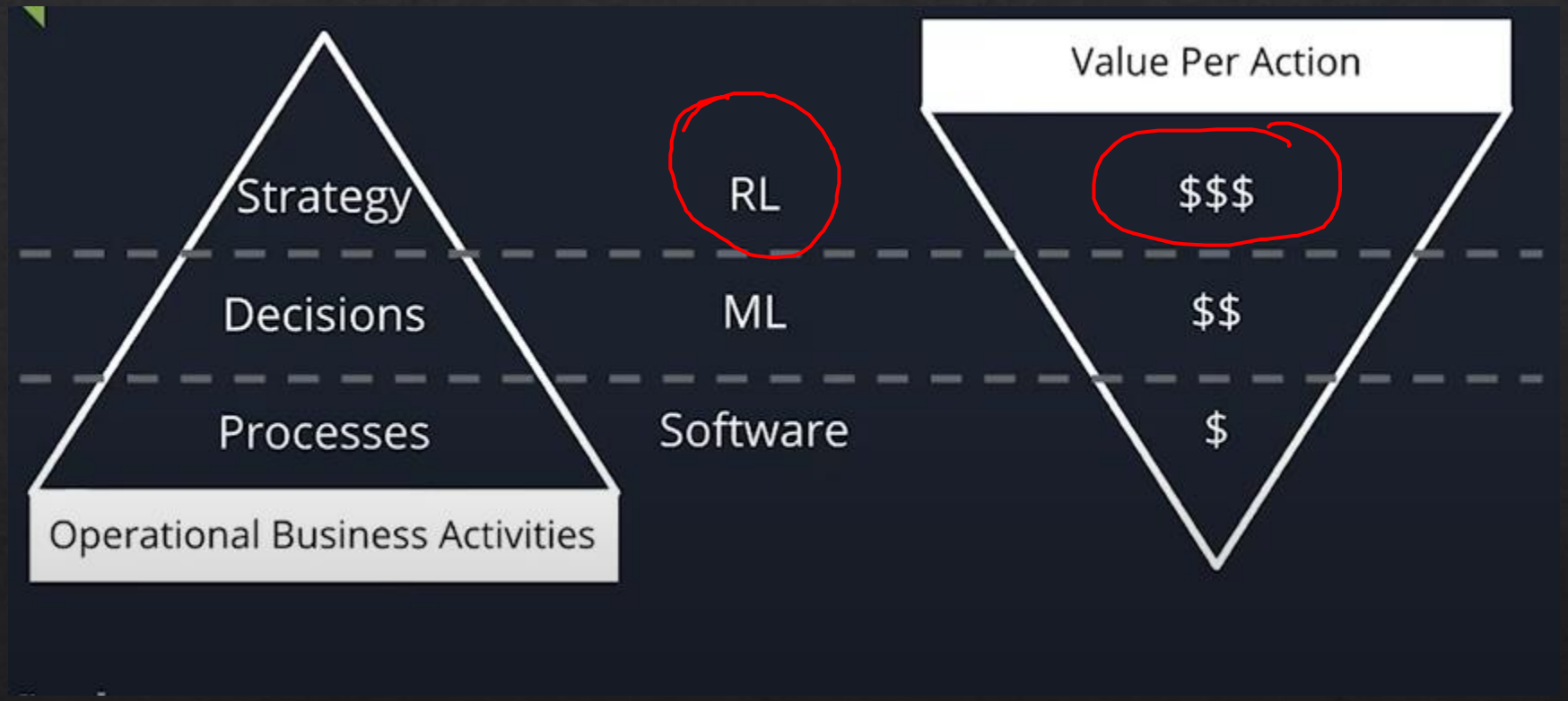


# RL en Negocios



# RL en Negocios


# RL en Negocios

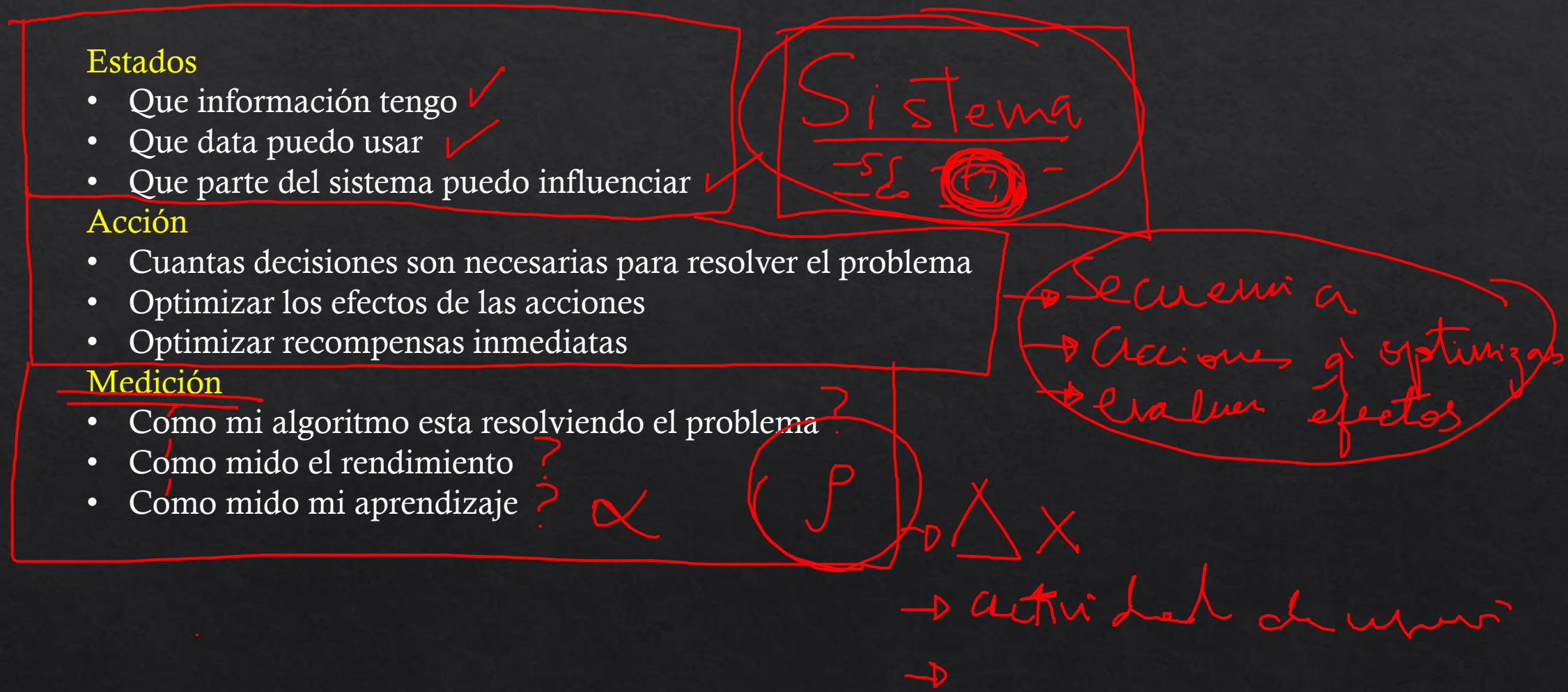


# RL en Toma de Decisiones

Estados
- Que información tengo
- Que data puedo usar
- Que parte del sistema puedo influenciar

Acción
- Cuantas decisiones son necesarias para resolver el problema
- Optimizar los efectos de las acciones
- Optimizar recompensas inmediatas

Medición
- Como mi algoritmo esta resolviendo el problema
- Como mido el rendimiento
- Como mido mi aprendizaje



# RL en Toma de Decisiones

Estados

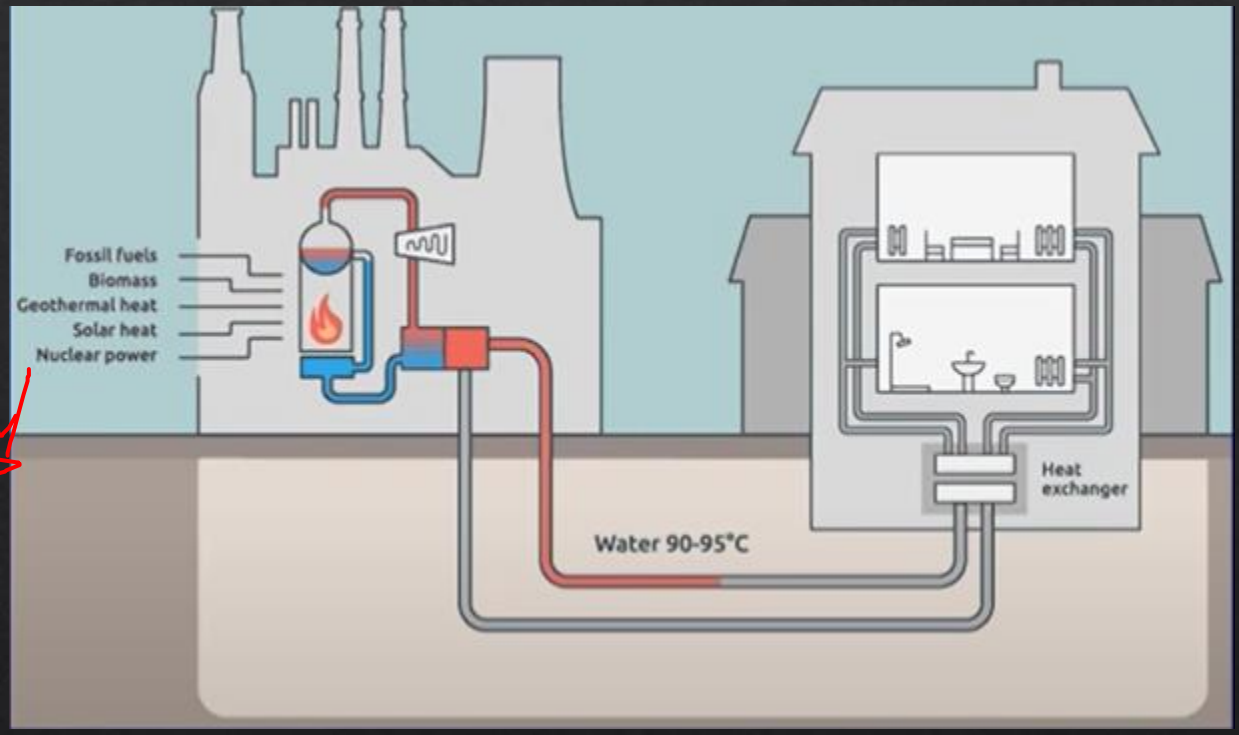
- Clima histórico
- Temperatura instantanea

Acción

Decidir la temperatura del agua deseada

Medición

Diferencia entre 22 °C y temperatura de la pieza



# RL en Toma de Decisiones

Estados
Costos
Acción
Precios del ticket diario
Medición
Maximización de utilidades diarios

# RL en Toma de Decisiones

Estados

Demanda histórica, inventario, ítems ordenados

Acción

Cobertura semanal

Medición

Cobertura promedio + margen por ítem

# RL en Toma de Decisiones

- Mejorar tiempo de excavación
  - Que acciones hacer para mejorar el tiempo (mantener, parar, retroceder)
  - ?