Reinforcement Learning Model in Automated Greenhouse Control

Ph.D. Francisco Javier Ferrández Pastor

PhD. Sara Alcañiz Lucas

MSc. Sofia Pardo

Ph.D. José María Cámara Zapata

Mag. José Antonio Brenes Carranza

15th International Conference on Ubiquitous Computing and Ambient Intelligence

November 28-30, 2023. Riviera Maya, México

About us

PhD. Francisco Javier Ferrández Pastor
Computer Scientist
DTIC, University of Alicante

PhD. José María Cámara Zapata
Agronomist
CIAGRO, University Miguel Hernández

MSc. Sofía Pardo Pina
Agronomist
CIAGRO, University Miguel Hernández

PhD. Sara Alcañiz Lucas
Greenhouse technician
PCA, University of Alicante

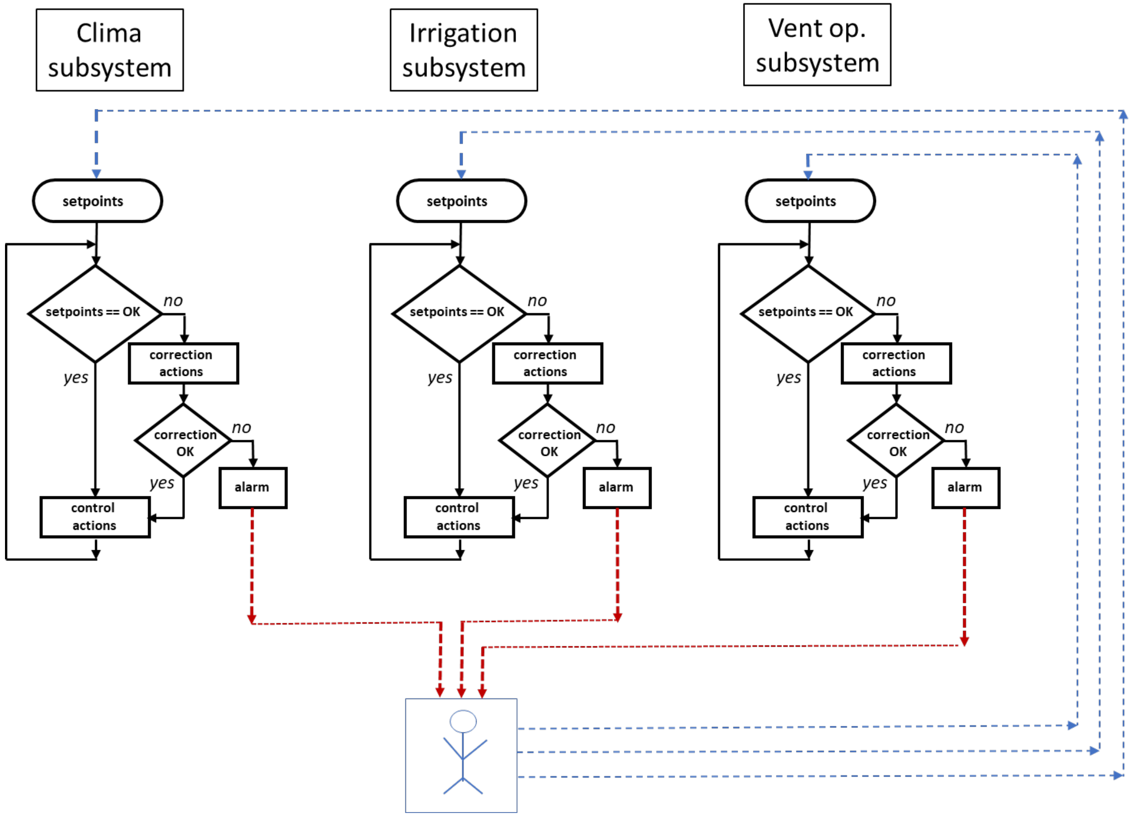
Mag. José Antonio Brenes Carranza
Computer Scientist
CITIC, University of Costa Rica

Motivation

- Why we work to support agriculture?

Accelerated population growing	Increased demand for food	Pollution of land and water	High costs of technology solutions
Need for technical and expert assistance	Potential of emergent technologies	Artificial intelligence boom	

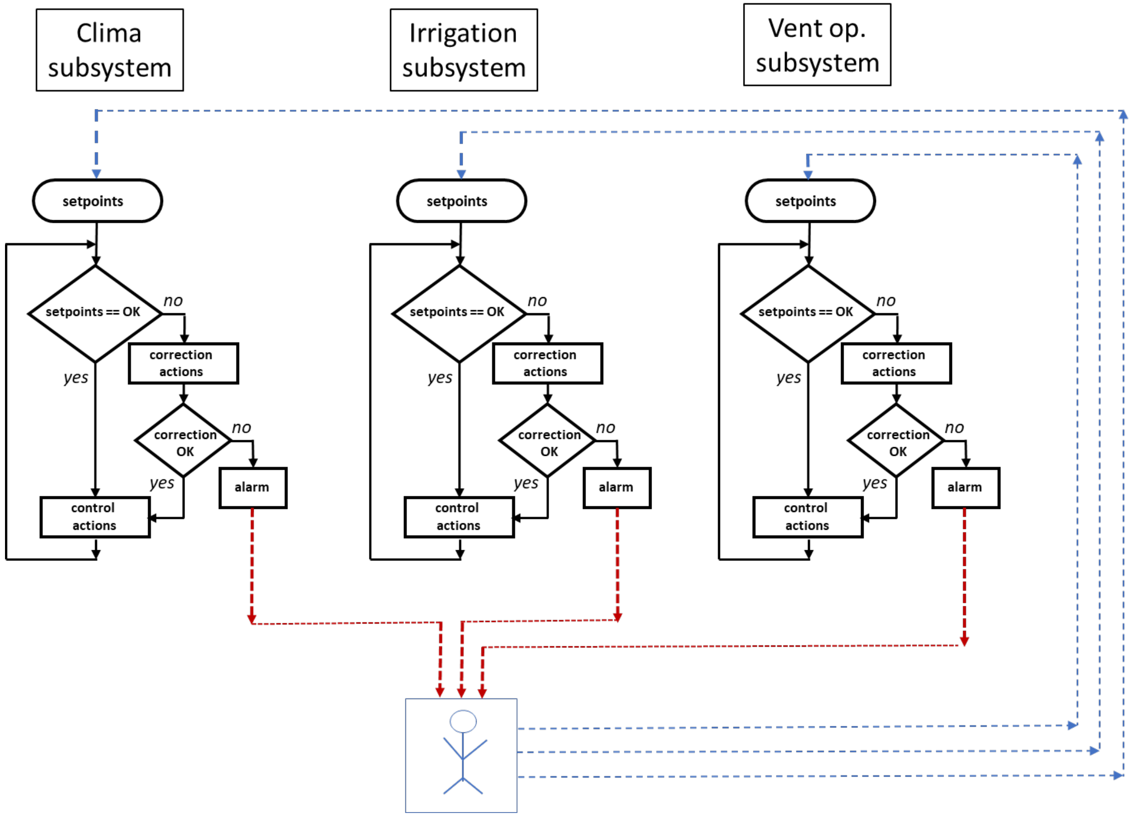
Background



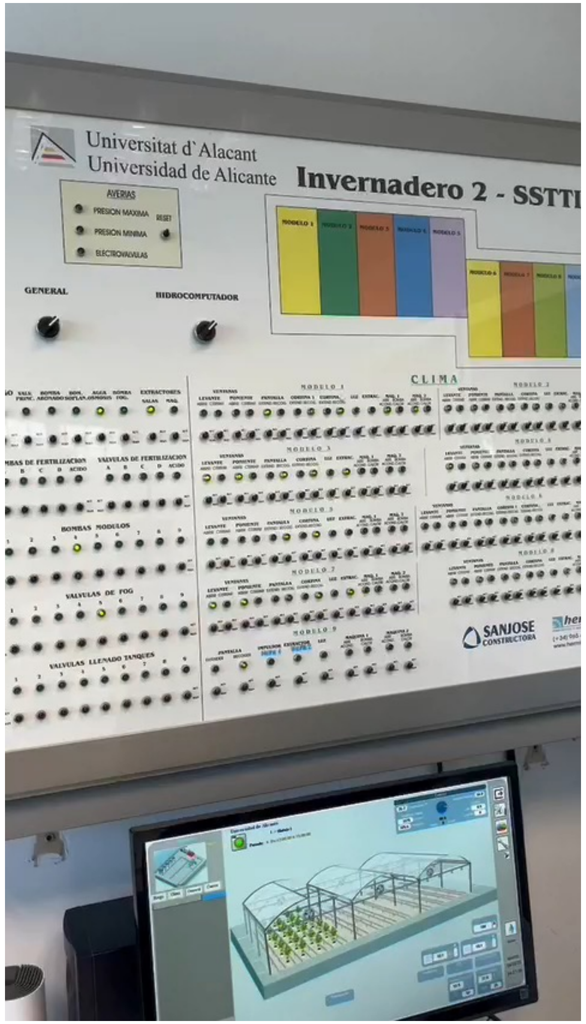
Automated scenario

Automated greenhouse

Background

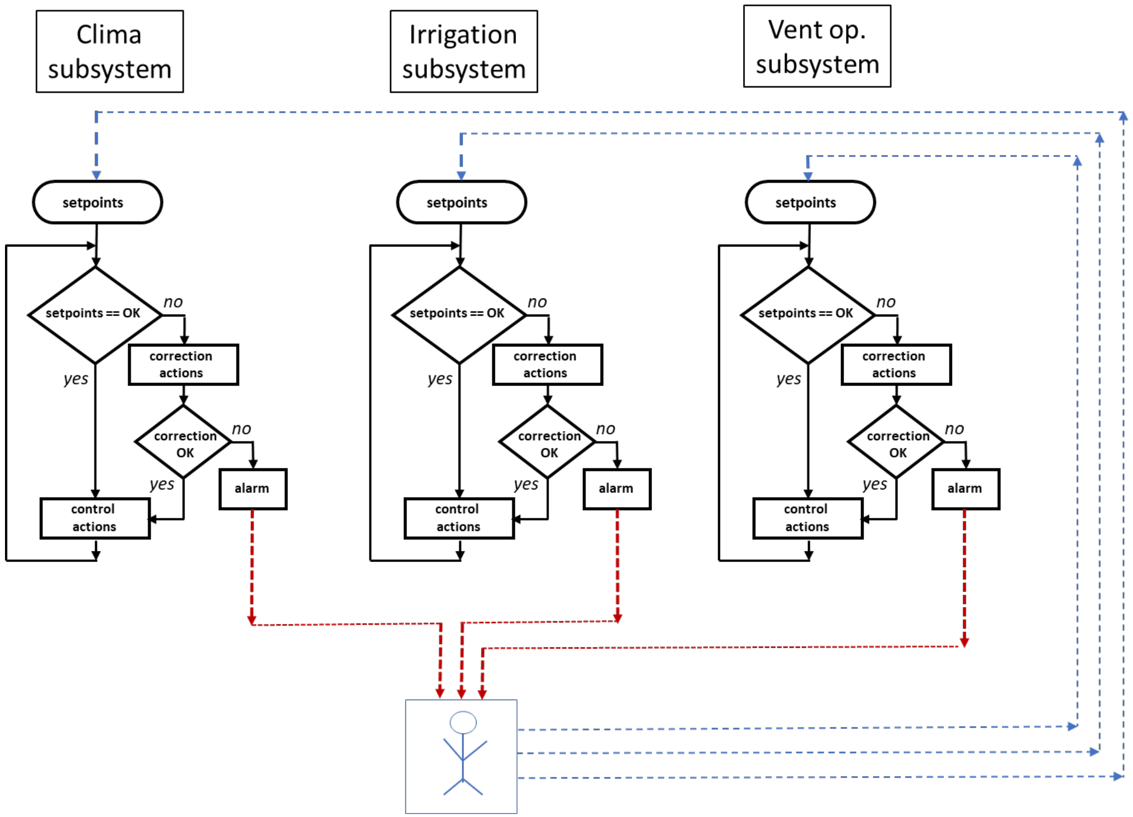


Automated scenario

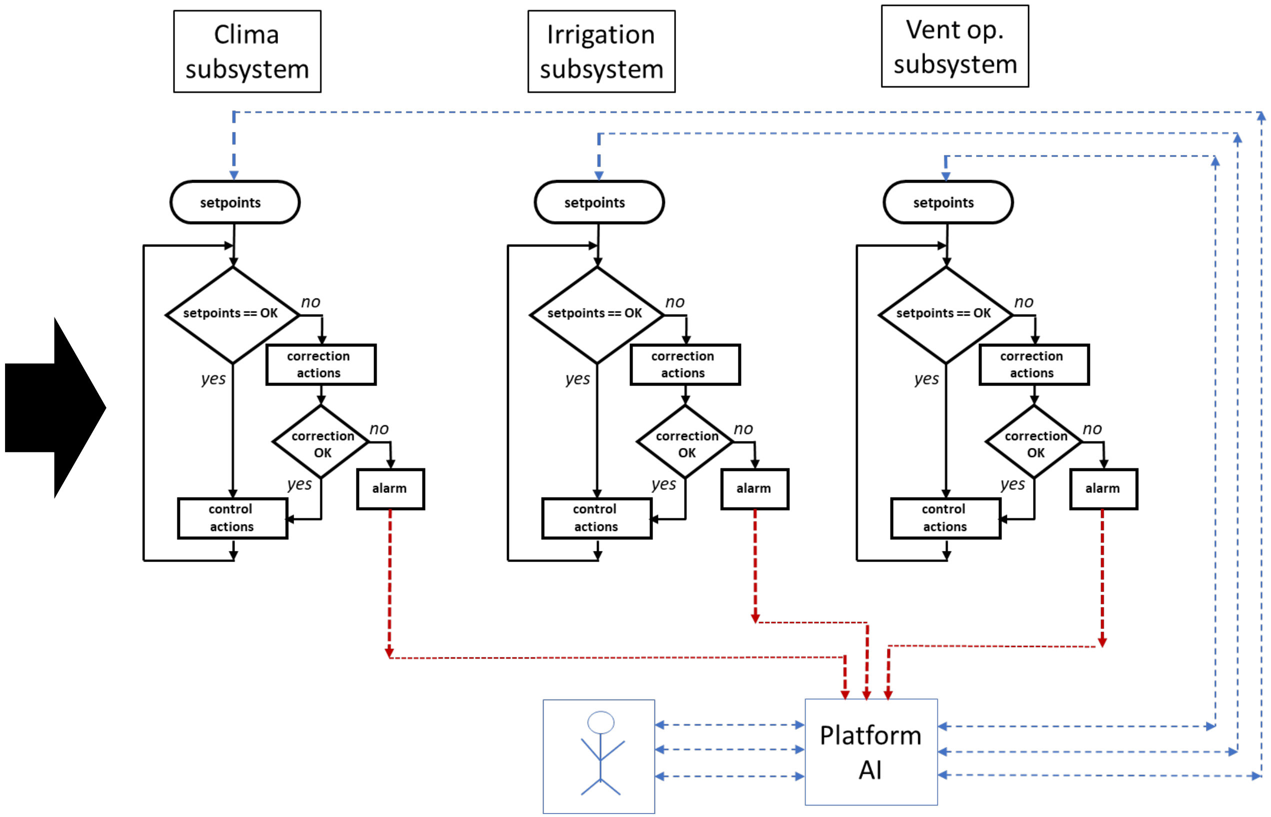


Greenhouse's main computer

Background



Automated scenario



Smart scenario

Some concepts

- Reinforcement Learning

Machine Learning technique

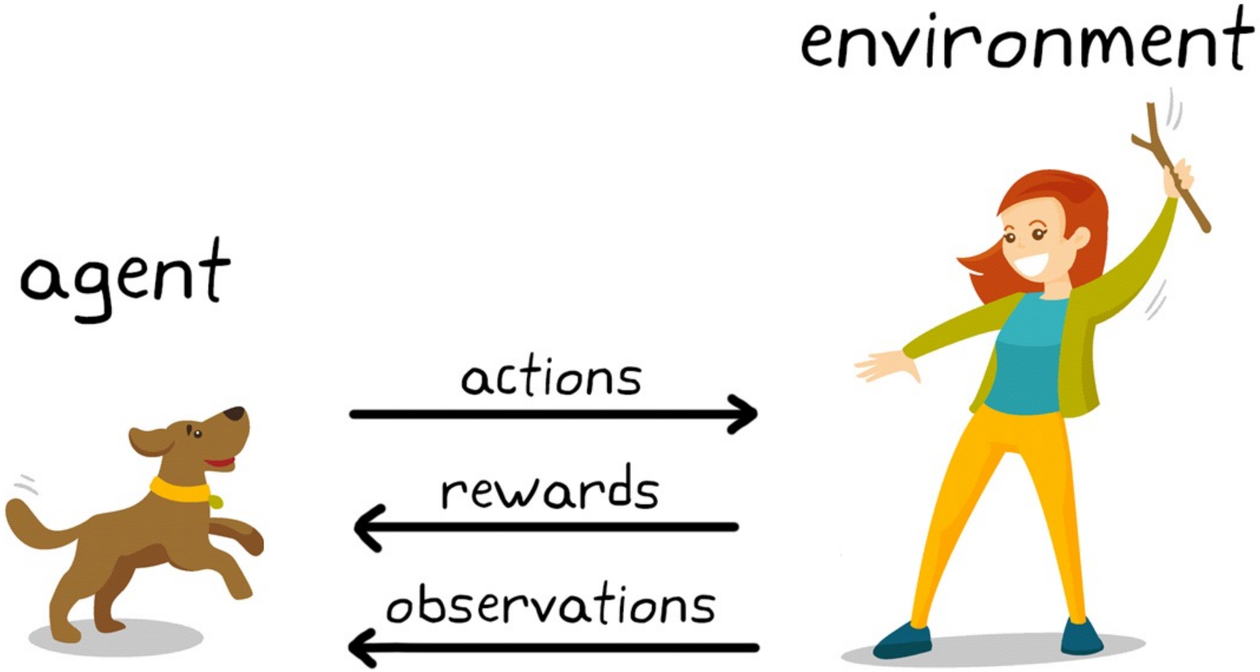
Computer agent

Trial-and-error interactions

Dynamic environment

Some concepts

- Reinforcement Learning – how it works?



Some concepts

• Q-Learning algorithm

Action	Fetching	Sitting	Running
Start	0	0	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	0	0
End	0	0	0

Action	Fetching	Sitting	Running
Start	0	1	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	0	0
End	0	0	0

Some concepts

- Q-Learning algorithm

Action	Fetching	Sitting	Running
Start	0	1	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	0	0
End	0	0	0

Action	Fetching	Sitting	Running
Start	0	1	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	34	0
End	0	0	0

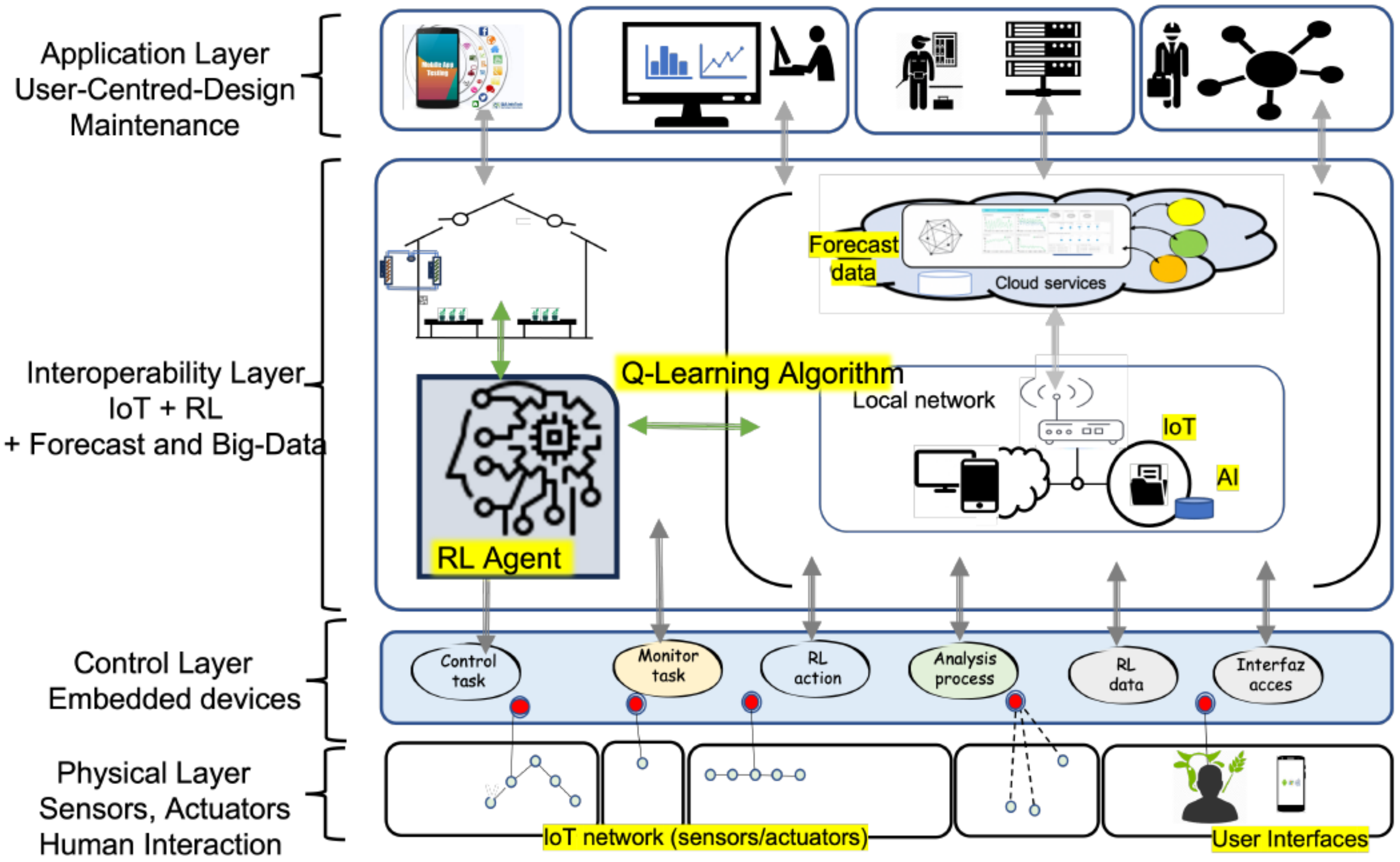
Some concepts

- Q-Learning algorithm

Action	Fetching	Sitting	Running
Start	0	1	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	34	0
End	0	0	0

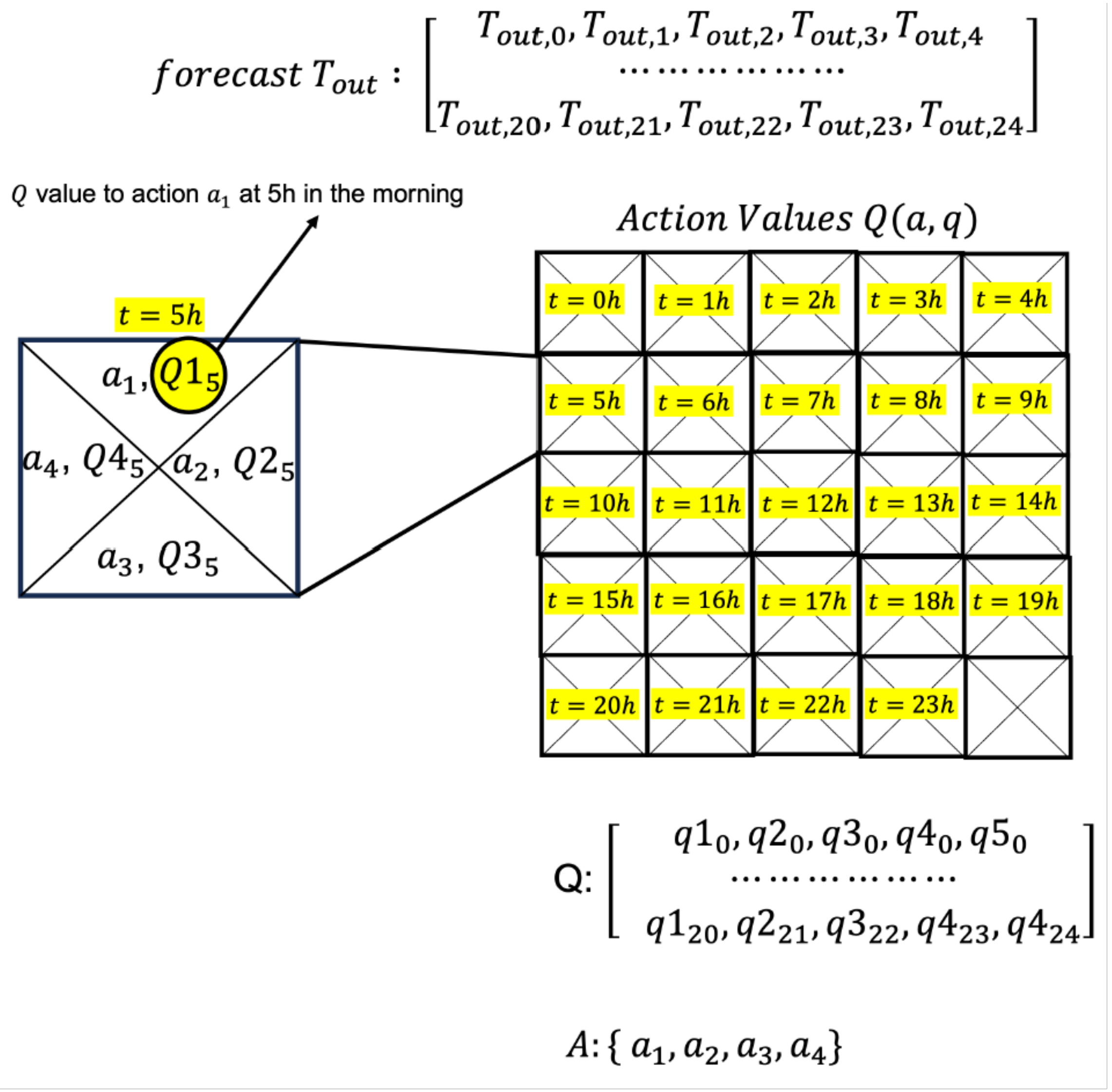
Action	Fetching	Sitting	Running
Start	5	7	10
Idle	2	5	3
Wrong Action	2	6	1
Correct Action	54	34	17
End	3	1	4

The proposed intelligent system



The algorithm

- The predictive model is based on the use of the prediction of the outside temperature. This allows calculations in the RL algorithm to predict the best control action to perform in the greenhouse thermal control system.



The algorithm

Algorithm 1 Q-learning: Learn function $Q : \mathcal{X} \times \mathcal{A} \rightarrow \mathbb{R}$

Require:

States $\mathcal{X} = \{1, \dots, n_x\}$

Actions $\mathcal{A} = \{1, \dots, n_a\}$, $\quad A : \mathcal{X} \Rightarrow \mathcal{A}$

Reward function $R : \mathcal{X} \times \mathcal{A} \rightarrow \mathbb{R}$

Learning rate $\alpha \in [0, 1]$, $\alpha = 1$

Discounting factor $\gamma \in [0, 1]$

procedure QLEARNING($\mathcal{X}$, A, R, T, α, γ)

  Initialize $Q : \mathcal{X} \times \mathcal{A} \rightarrow \mathbb{R}$ arbitrarily

  while Q is not converged **do**

    Start in state $s \in \mathcal{X}$

    while s is not terminal **do**

      Calculate π according to Q and exploration strategy

      $a \leftarrow \pi(s)$

      $r \leftarrow R(s, a)$ ▷ Receive the reward

      $s' \leftarrow T(s, a)$ ▷ Receive the new state

      $Q(s', a) \leftarrow Q(s, a) + \alpha \cdot (r + \gamma \cdot \max_{a'} Q(s', a'))$

      $s \leftarrow s'$

  return Q

Results

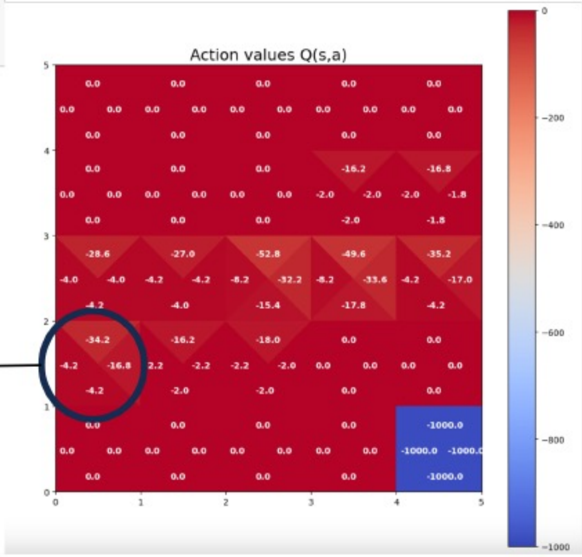
```
In [2]: for i in range(24):
            hora=i*60
            q_learning(action_values,exploratory_policy,target_policy,1000,hora)
        print(action_values)
[[[    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]]

 [[    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [  -16.2     -2.      -2.      -2. ]
  [  -16.8     -1.8     -1.8     -2. ]]

 [[  -28.6     -4.      -4.2     -4. ]
  [  -27.      -4.2     -4.      -4.2]
  [  -52.8    -32.2    -15.4     -8.2]
  [  -49.6    -33.6    -17.8     -8.2]
  [  -35.2    -17.      -4.2     -4.2]]

 [[  -34.2    -16.8     -4.2     -4.2]
  [  -16.2     -2.2     -2.      -2.2]
  [  -18.      -2.      -2.      -2.2]
  [    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]]

 [[    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [    0.       0.       0.       0. ]
  [-1000.   -1000.   -1000.   -1000. ]]]
```



at 15:00 h

$$\begin{bmatrix} Q1 = -34.2\,, a_1 \rightarrow T_{setpoint} = (T_{max}-1) \\ Q2 = -16.8, a_2 \rightarrow T_{setpoint} = (T_{max}-2) \\ Q3 = -4.2\,, a_3 \rightarrow T_{setpoint} = (T_{max}-3) \\ Q4 = -4.2\,, a_4 \rightarrow T_{setpoint} = (T_{max}-4) \end{bmatrix}$$

$$T_{setpoint} = (T_{max}-4)$$

Conclusions

- We propose a novel approach to greenhouse control by leveraging a single-agent RL centralized controller
- The results show that the RL-based approach can effectively automate and optimize set-point selection in greenhouse control systems
- As the agricultural industry seeks more efficient and sustainable solutions, the insights from this study contribute valuable knowledge towards the development of intelligent and autonomous greenhouse control systems
- The RL model forms part of an integral smart farming platform that monitors and controls greenhouse conditions in a sustainable way, considering multiple subsystems and controllers

Future work

- This work is associated with in-progress doctoral research on developing a technological architecture to guide smart farming application development and deployment.
- RL algorithms often require a learning period during which the system may not perform optimally. We porpose to capture a large number of data on the real behavior of the system without the use of RL, by using IoT devices.
- Future work may focus on scaling up the approach for multi-agent systems and further exploring the RL algorithm's performance in various agricultural settings.

Funding & Acknowledgements

- Funding → This study is part of the AGROALNEXT program (AGROALNEXT/2022/048) and has been supported by MCIN with funding from the European Union NextGenerationEU (PRTR-C17.I1) and the Generalitat Valenciana. This study was partially supported by the Research Center for Communication and Information Technologies (CITIC) and the School of Computer Science and Informatics (ECCI) at the University of Costa Rica, Research Project No. 834-B9-189.

- Acknowledgements → We are very grateful to Palmeera Farms (Palmeera) Biotechnology-based company, member of the Spanish Association of Biotechnology Companies (Asebio) and attached to the Alicante Science Park (PCA) for their collaboration in this work.

Universitat d'Alacant

UNIVERSIDAD DE COSTA RICA

UNIVERSITAS Miguel Hernández

Thank you for your attention!

Mag. José Antonio Brenes Carranza

joseantonio.brenes@ucr.ac.cr

PhD. Francisco Javier Ferrández Pastor

fjferran@dtic.ua.es