

# Reinforcement Learning Quarto

Norah Jones

Invalid Date

## **Table of contents**

# Preface

This is a Quarto book.

To learn more about Quarto books visit <https://quarto.org/docs/books>.

1 + 1

[1] 2

# 1 Introduction

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

```
1 + 1
```

```
[1] 2
```

## 2 Activate the Core Packages

```
library(bibtex)
library(tidyverse) ## Brings in a core of useful functions
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.5.1      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr      1.0.2
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(gt)          ## Tables
## Specific packages
library(milestones)
## Initialize defaults
## Initialize defaults
column <- lolli_styles()

data <- read_csv(col_names=TRUE, show_col_types=FALSE, file='rl_time_line.csv')
```

```
## Sort the table by date
data <- data |>
  arrange(date)

## Build a table
gt(data) |>
  #cols_hide(columns = event) |>
  tab_style(cell_text(v_align = "top"),
            locations = cells_body(columns = date)) |>
  tab_source_note(source_note = "Source: Sutton and Barto (2018)")
```

| date | event  | referen |
|------|--|---------|
| 1911 | Primer idea del "trial-and-error learning" (TaEL)                                    | NA      |
| 1927 | Aparece el termino "Reinforcement" en el contexto del aprendizaje animal             | NA      |
| 1948 | Turing describe un diseno para un sistema "pleasure-pain system"                     | NA      |
| 1954 | Minsky Farley y Clark publican sus investigaciones sobre TaEL.                       | NA      |
| 1957 | Aparece "Dynamic Programming" (DP)   | mundo   |
| 1957 | Introducen los Markov Decision Processes (MDP)                                       | NA      |
| 1959 | Comienza a utilizarse el "optimal control" (OC)                                      | paso    |
| 1959 | Se desarrolla extensamente la DP   | NA      |
| 1960 | Aparece el metodo de iteracion de politicas  | NA      |
| 1960 | Se utilizan por primera vez en ingeniería "Reinforcement" y "Reinforcement Learning" | NA      |
| 1960 | Se originan los "Learning Automata"  | NA      |
| 1961 | Se publica "Steps Toward Artificial Intelligence"                                    | NA      |
| 1961 | Se describe un sistema "TaEL" para el tic-tac-toe                                    | NA      |
| 1963 | Se desarrolla STeLLA   | NA      |
| 1972 | Trabajo de Klopff  | NA      |
| 1973 | Widrow, Gupta y Maitra modifican el LMS.   | NA      |
| 1973 | Teoría del aprendizaje de Bush y Mostelle  | NA      |
| 1977 | Werbos conecta el OC y PD  | NA      |
| 1978 | Sutton desarrolla las ideas de Klopff  | NA      |
| 1986 | Introducen los clasificadores  | NA      |
| 1989 | Watkins integra los metodos de aprendizaje (MA)                                      | NA      |
| 1996 | Aparece el termino "Neurodynamic Programming"  | NA      |
| 2003 | "Reinforcement Learning" en economí  | NA      |
| 2012 | Vision general del Reinforcement Learning y Juegos                                   | NA      |
| NA   | NA   | NA      |

Source: Sutton and Barto (2018)

```
## Adjust some defaults
column$color <- "orange"
column$size <- 15
column$source_info <- "Source: Sutton and Barto (2018)"

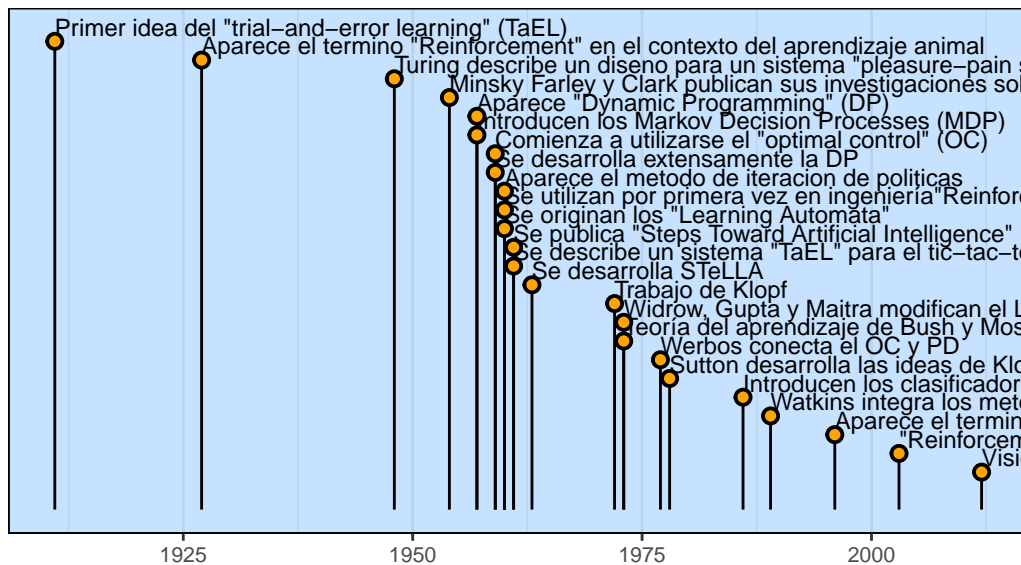
## Milestones timeline
milestones(datatable = data, styles = column)
```

Warning: Removed 1 row containing missing values or values outside the scale range (`geom\_text()`).

Warning: Removed 1 row containing missing values or values outside the scale range (``geom_segment()``).

Warning: Removed 1 row containing missing values or values outside the scale range (``geom_point()``).

Warning: Removed 25 rows containing missing values or values outside the scale range (``geom_segment()``).



Source: Sutton and Barto (2018)

## 2.1 EJERCICIO 1

En el aprendizaje reforzado un agente aprende a tomar decisiones (acciones) a través de la interacción con su entorno y recibiendo recompensas o castigos en función de las mismas, a diferencia del aprendizaje supervisado, ya que en este tipo de aprendizaje automático, un modelo se entrena utilizando un conjunto de datos que incluye tanto las entradas como las salidas correspondientes (etiquetas). es decir, consiste en aprender a partir de un conjunto de ejemplos ya etiquetados y proporcionados por un supervisor externo con conocimientos. Por lo que en este tipo de aprendizaje Cada ejemplo describe una situación específica, y además existe una etiqueta que indica la acción adecuada que el sistema debe tomar en esa situación, El objetivo de este tipo de aprendizaje es que el sistema generalice sus respuestas para que actúe correctamente en situaciones que no están