DesignCon® 2019
WHERE THE CHIP MEETS THE BOARD

Spec-Driven CTLE Model Synthesis through Reinforcement Learning

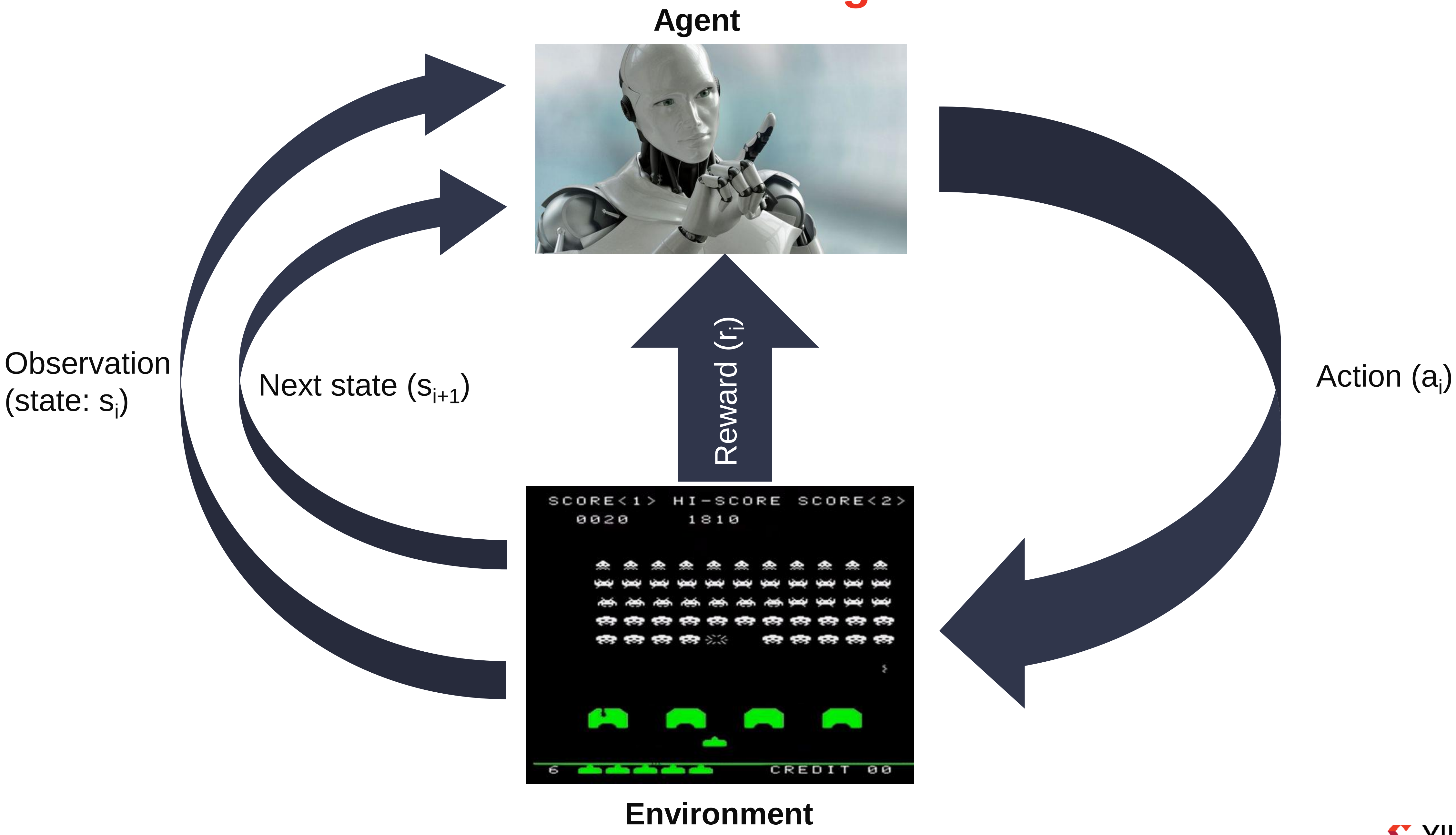
.com

Presenter: Daniel Wu, danielw@xilinx.com

Agenda

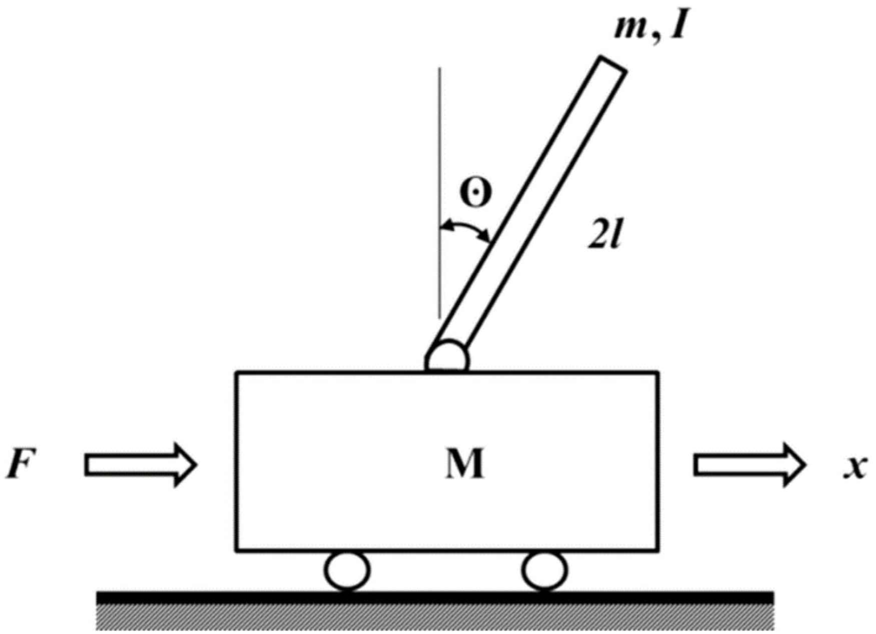
- Introduction to Reinforcement Learning
  - Application Examples
- Synthesis of CTLE Modeling
- Spec-Driven w/ Reinforcement Learning
- Examples
- Summary

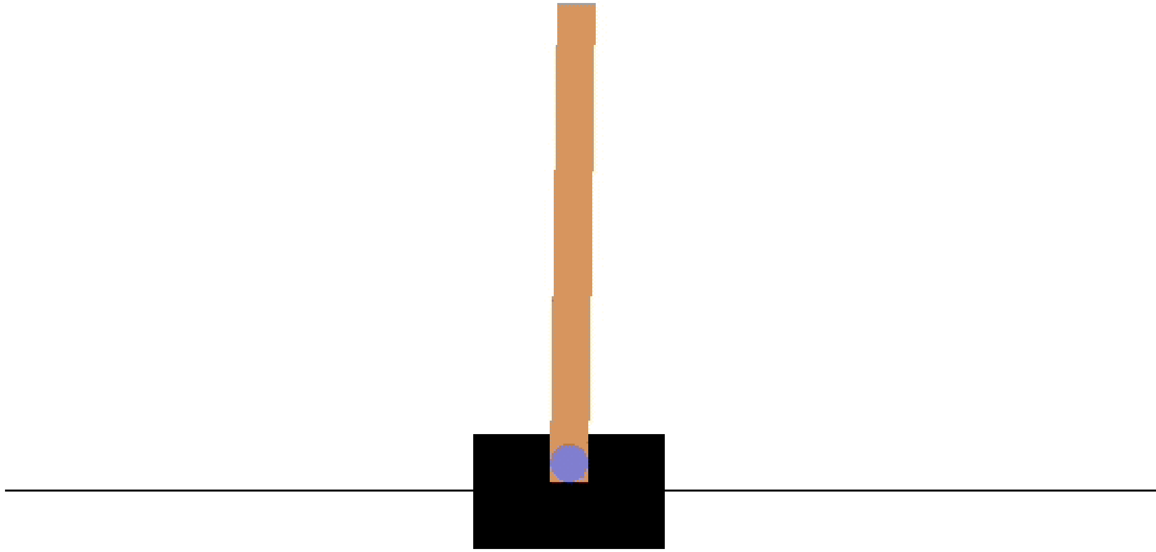
Introduction of Enforcement Learning



Introduction of Enforcement Learning

Application Example: Cart-Pole Problem





Objective: Balance a pole on top of a movable cart

State (Observation): angle, angular speed, position, horizontal velocity

https://keon.io/deep-q-learning/

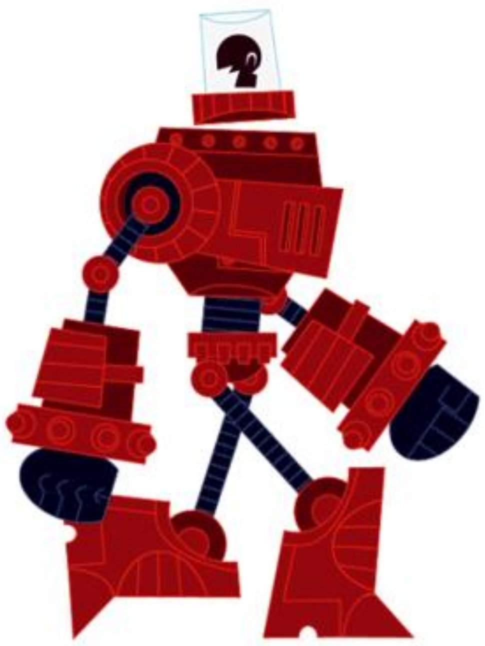
Action: horizontal force applied on the cart

Reward: +1 at each time step if the pole is upright

http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture14.pdf

Introduction of Enforcement Learning

Application Example: Robot Locomotion



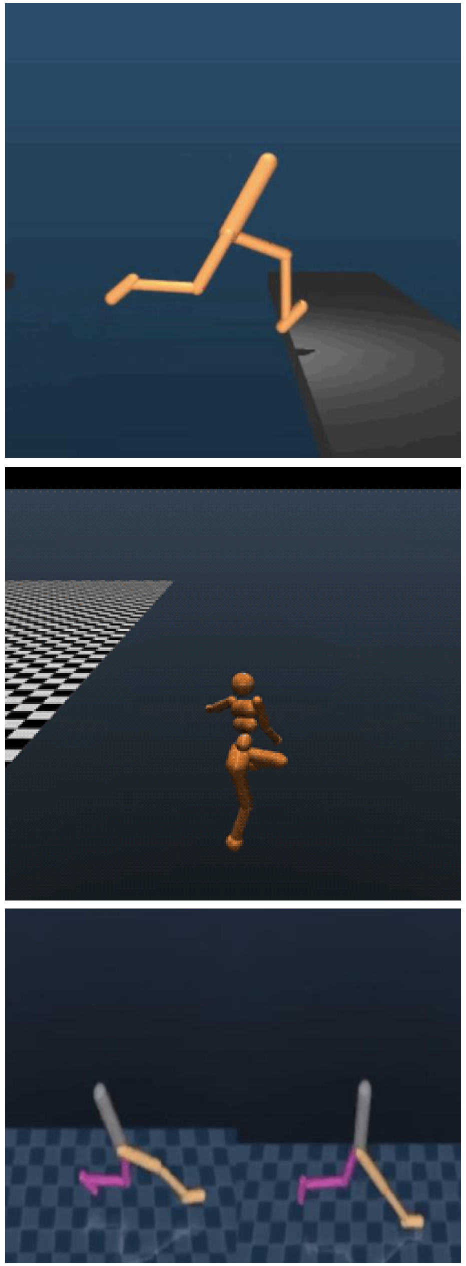
Objective: Make the robot move forward

State (Observation): Angle and position of the joints

Action: Torques applied on joints

Reward: +1 at each time step upright & forward movement

http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture14.pdf



Google Deepmind Ai

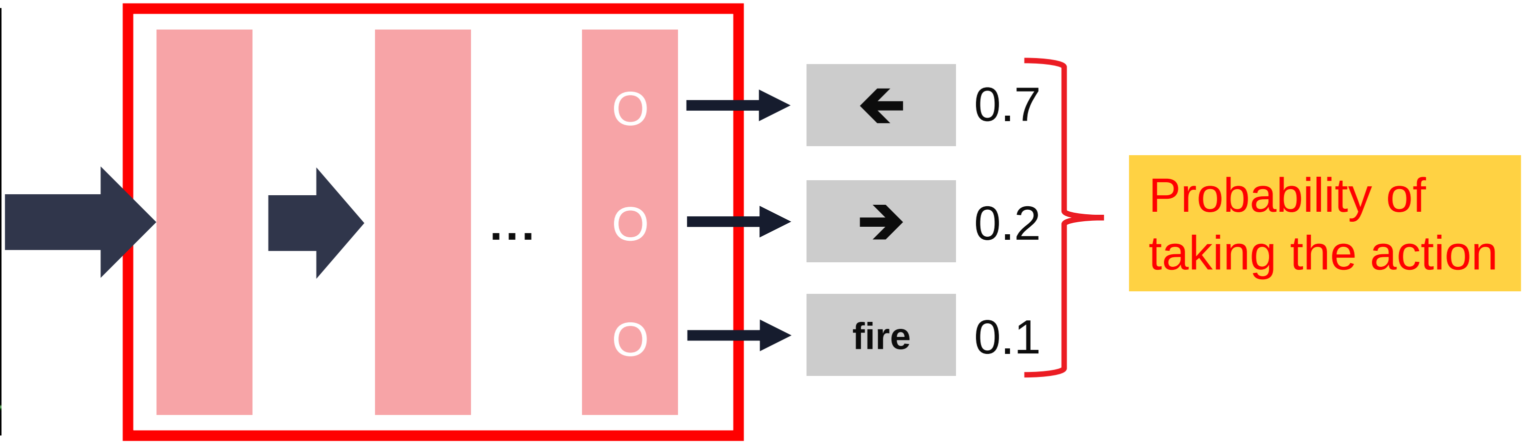
Introduction of Enforcement Learning

Application Examples: Atari, Space Invader, and etc.

Pixels

or

NN as actor



26 Feb 2015, vol 518

Objective: Complete the game with highest score

State (Observation): Raw pixel inputs of the game state

Action: Game controls e.g. left, right, and/or fire

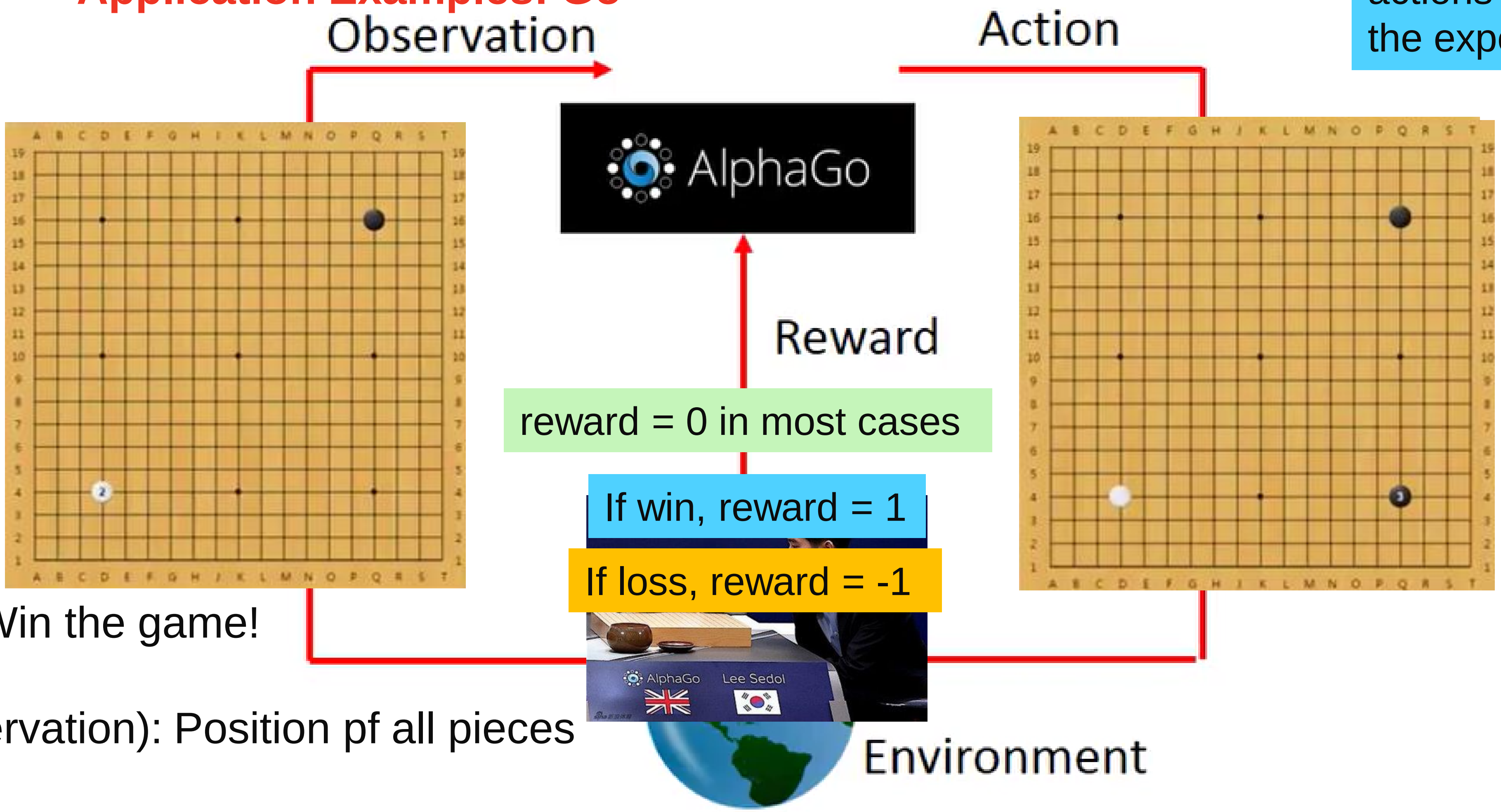
Reward: Score increase at each time step

http://speech.ee.ntu.edu.tw/~tlkagk/courses_ML17_2.html

XILINX

Introduction of Enforcement Learning
Application Examples: Go

Agent learn to take actions to maximize the expected reward.



28 Jan 2016, vol 529

Objective: Win the game!

State (Observation): Position pf all pieces

Action: Where to put the next piece down

Reward: 1 if win at the end of the game, 0 otherwise

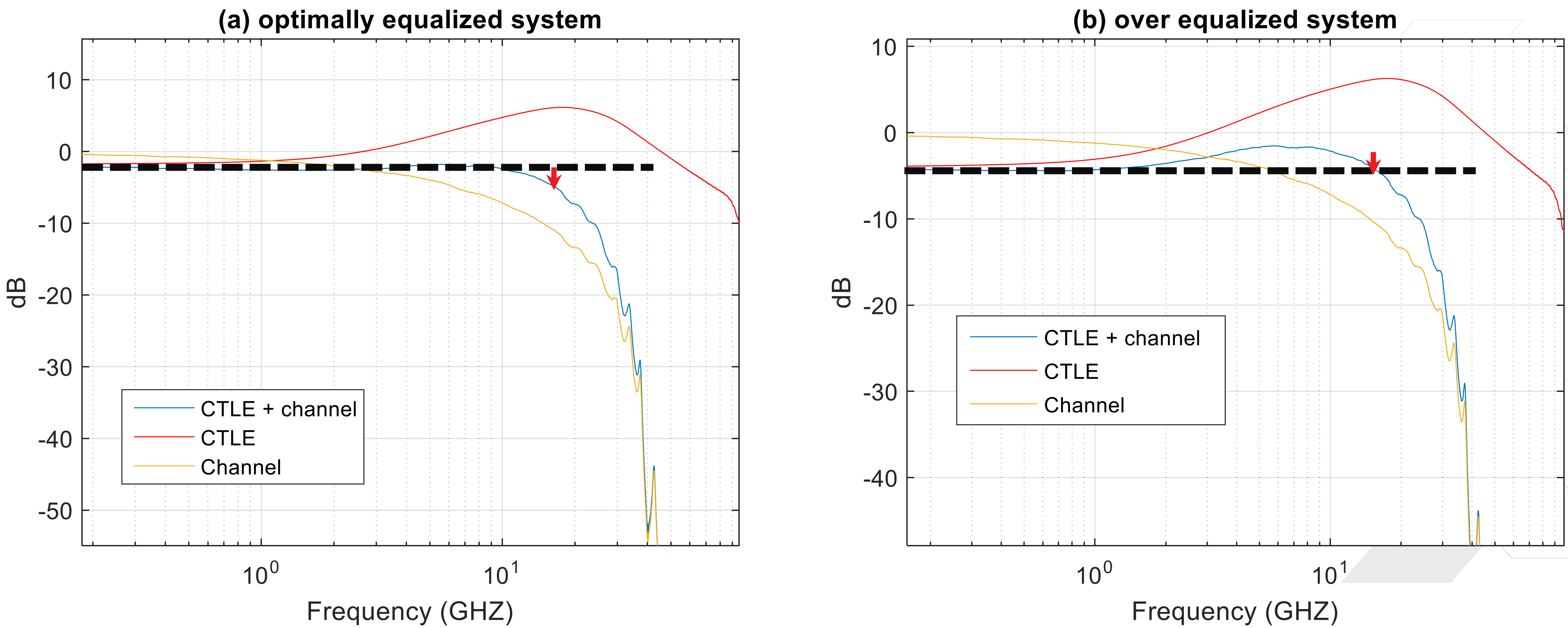
http://speech.ee.ntu.edu.tw/~tlkagk/courses_ML17_2.html

XILINX

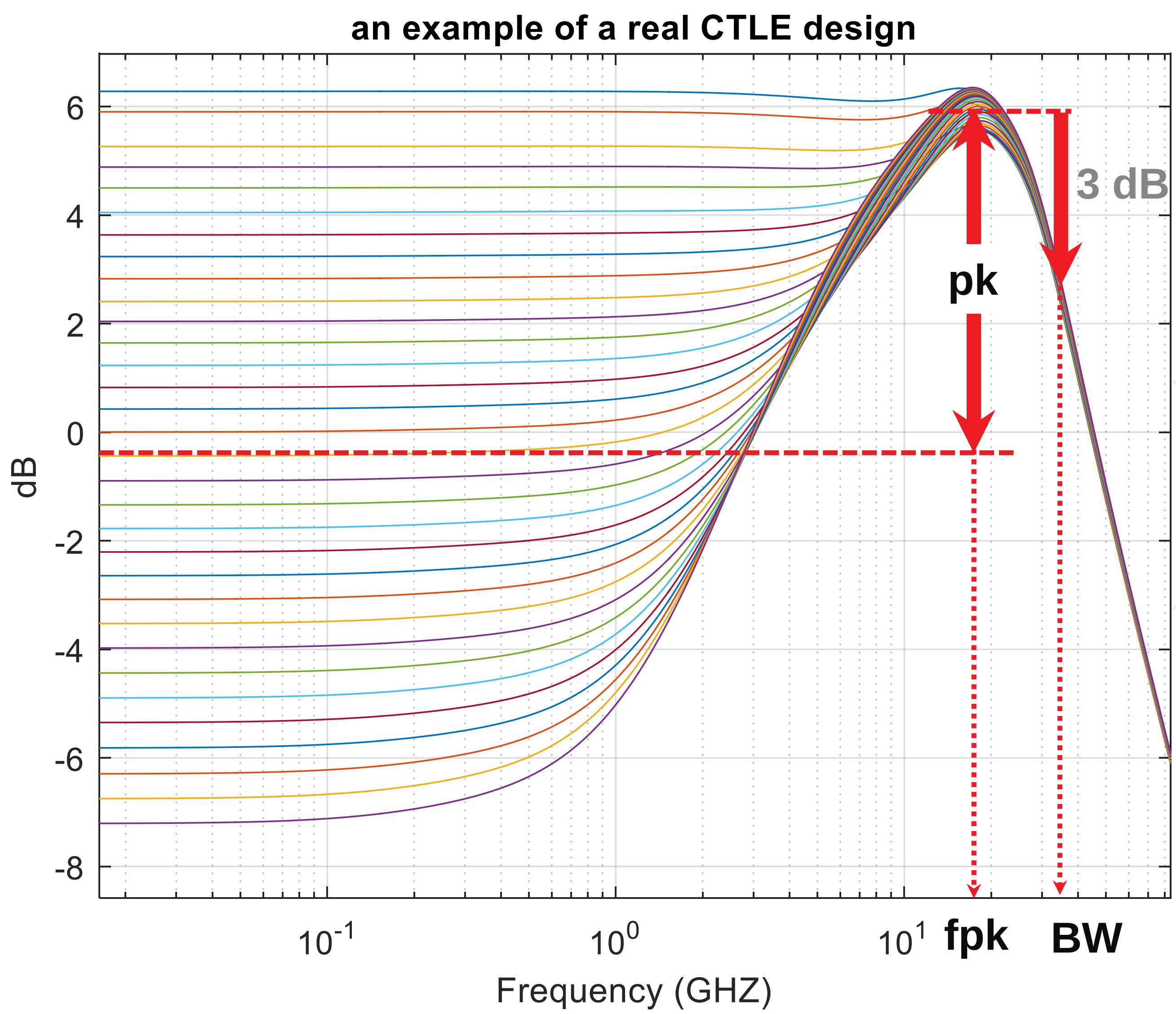
Agenda

- Introduction to Reinforcement Learning
  - Application Examples
- **Synthesis of CTLE Modeling**
- Spec-Driven w/ Reinforcement Learning
- Examples
- Summary

Synthesis of CTLE Model



CTLE Spec


an example of a real CTLE design
6
4
2
0
-2
-4
-6
-8
dB
10^{-1}
10^{0}
10^{1}
fpk
BW
Frequency (GHZ)
3 dB
pk

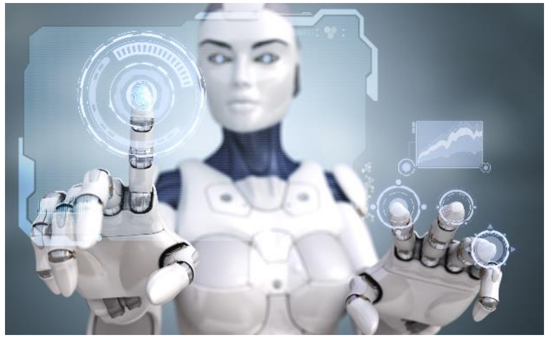
Agenda

- Introduction to Reinforcement Learning
  - Application Examples
- Synthesis of CTLE Modeling
- **Spec-Driven w/ Reinforcement Learning**
- Examples
- Summary

Spec-driven Modeling w/ Reinforcement Learning

state: s_i
[(s_fpk$_i$, s_pk$_i$), s_bw$_i$]

Agent



Next state: s_{i+1}
[(s_fpk$_{i+1}$, s_pk$_{i+1}$), s_bw$_{i+1}$]

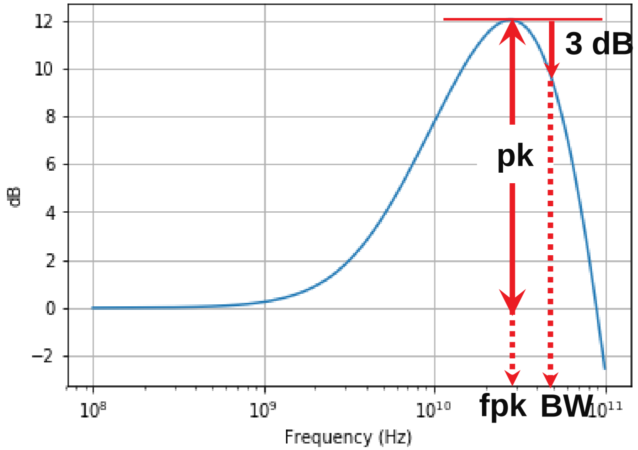
Reward (r_i)

$$Dist_i = \left(f\_pk_i - f\_pk_{spec}\right)^2 + \beta_i\left(pk_i - pk_{spec}\right)^2$$

$$\begin{cases} Dist_i < Dist_{i-1} : \boldsymbol{reward} = +2 \\ else \quad : \boldsymbol{reward} = -2 \end{cases}$$

$$\begin{cases} \left|f\_pk_i - f\_pk_{spec}\right| < tol : \beta_i = \beta_{i-1}/2 \\ else \quad : \beta_i = 2\beta_i \end{cases}$$

$$\begin{cases} BW_i < BW_{spec} : \boldsymbol{reward} = +1 \\ else \quad : \boldsymbol{reward} = -1 \end{cases}$$



Environment

Action: a_i

```
[incZ1,      decZ1,     incP1,    incP1,    incP1, incP2,
 incZ1P1,    decZ1P1,   incZ2P1,  decZ2P2,  incP1P2, decP1P2,
 incZ1P1P2,  decZ1P1P2,
 mvUP,       mvDN,      mvPUP,    mvPDN,
 incZ1decP2, incP1decZ1]
```


XILINX

Bellman Equation

$$Q(s_t, a_t) = r(s_t) + \gamma \max(Q(s_{t+1}, a_{t+1}))$$

r: reward of each action

Q: Total reward of full course

a: Action

s: State

γ : Decay factor

Objective: Tune the CTLE performance into user-defined specs

State (Observation): CTLE's [(f_{pk}, pk), BW]

Action: Maximized the total reward (Q) from the action table of each iteration

Reward: +3 or +1 if the distance to specs are shorter, vice versa.

Agenda

- Introduction to Reinforcement Learning
  - Application Examples
- Synthesis of CTLE Modeling
- Spec-Driven w/ Reinforcement Learning
- **Examples**
- Summary

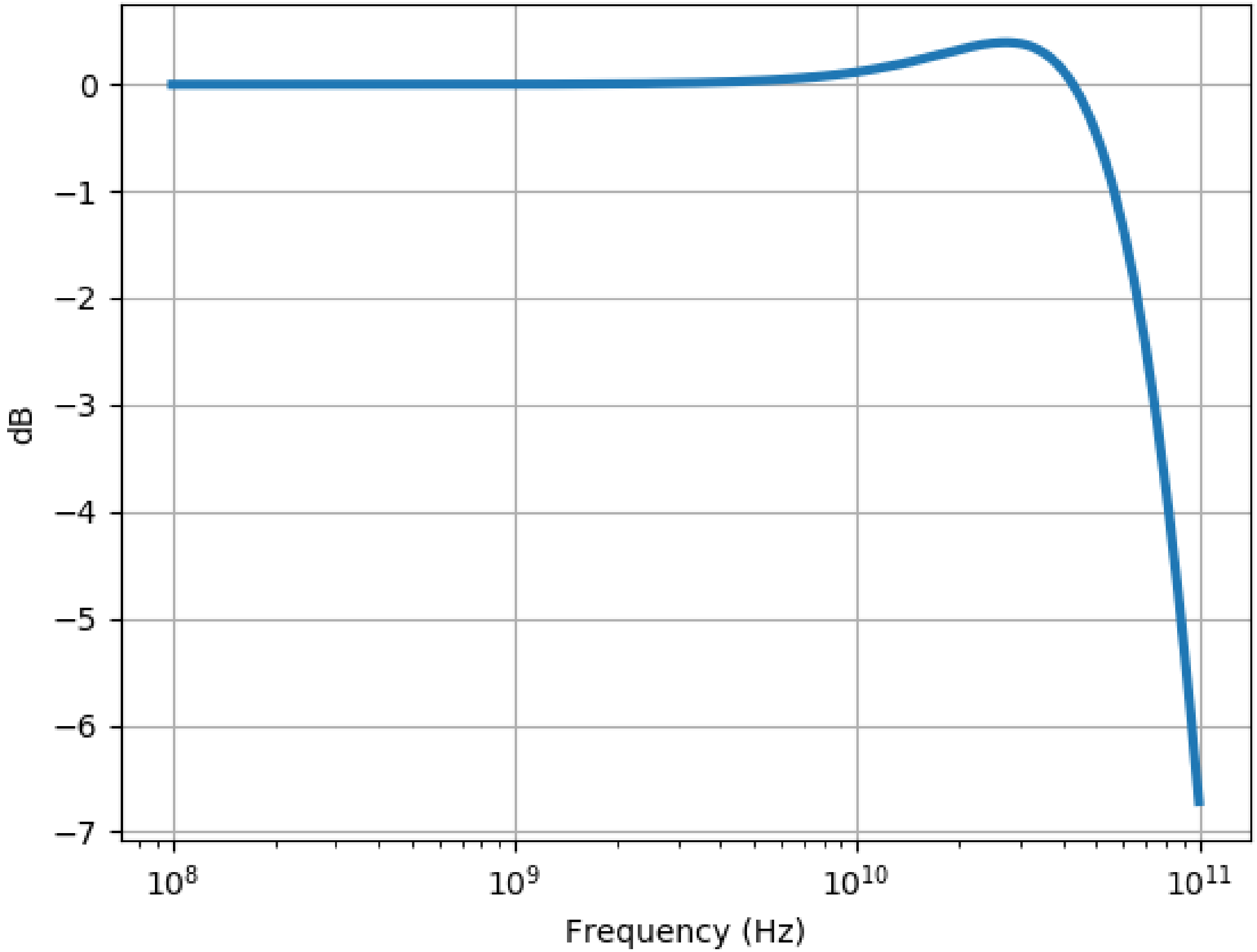
Examples: what if brute force?

```
for p1 in np.arange(1e9, 50e9, 3e8): # 164
    for p2 in np.arange(10e9, 100e9, 3e8): # 300
        for p3 in np.arange(10e9, 100e9, 3e8): # 300
            for p4 in np.arange(10e9, 100e9, 3e8): # 300
                for p5 in np.arange(10e9, 100e9,3e8): # 300
                    for z in np.arange(1e9, 10e9, 3e8):  # 30
                        freq, Hs = evalH([z], [2*np.pi*p1, 2*np.pi*p2, 2*np.pi*p3, 2*np.pi*p4, 2*np.pi*p5], 1)
                        (fpk, pk), bw = findpeak_bw(freq, Hs)
                        Xc = np.abs(fpk - peak_freq)/1e9
                        Yc = np.abs(pk - peak)
                        current_dist = Xc*Xc +    Yc*Yc
                        if current_dist < dist:
                            dist = current_dist
                            opt_pz = [z, [p1, p2, p3, p4, p5]]
```

164*300*300*300*300*30 ~ 40 Tera cases!

XILINX.

Examples: EL with a small $pk_{spec} = 0.1\ dB$



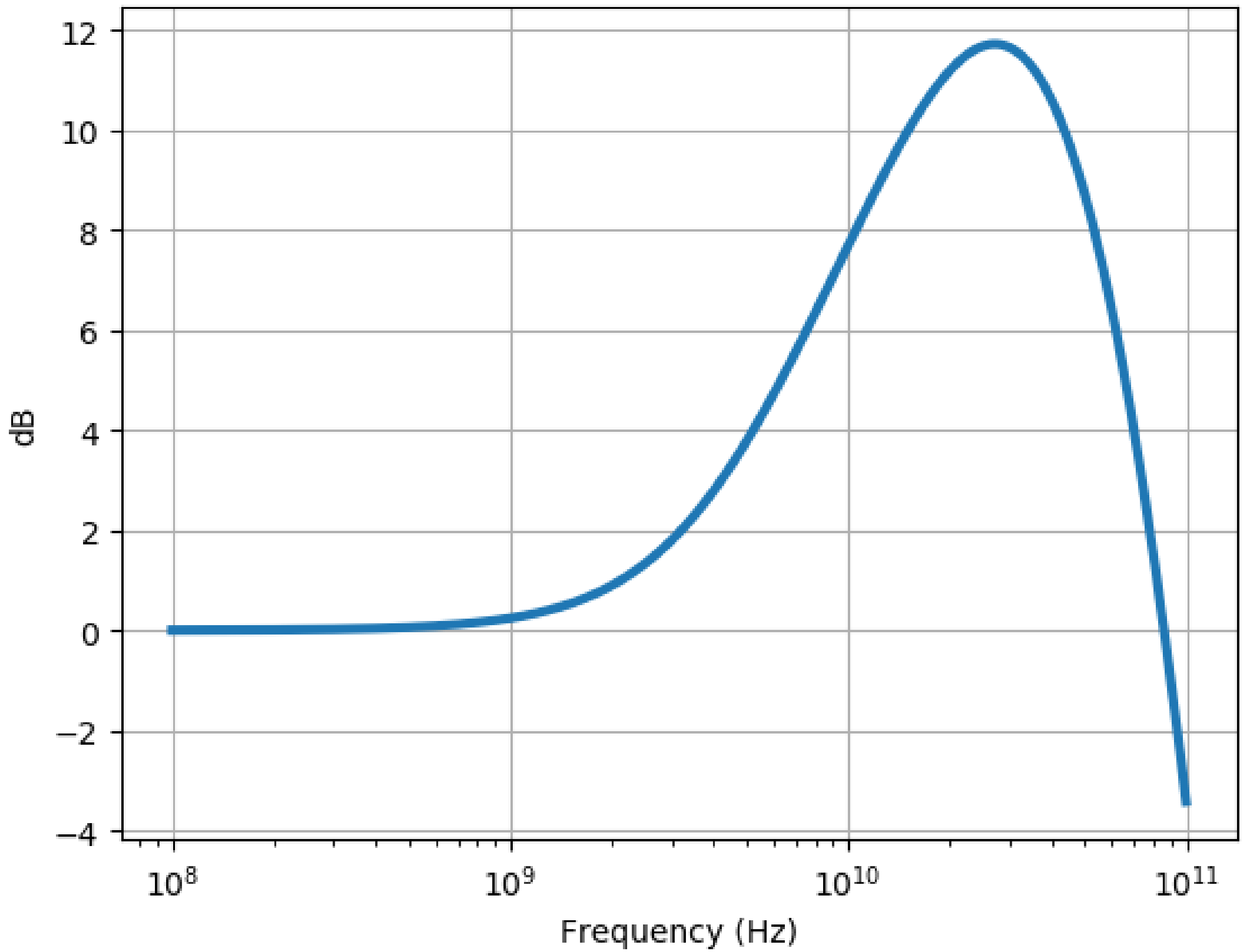
Number of Iteration = 1567

$[(f_{pk}, pk), BW]$ = [(27.3e9, 0.39), 71.1e9]

Elapsed time is 73 sec

XILINX

Examples: EL with large $pk_{spec} = 12\ dB$

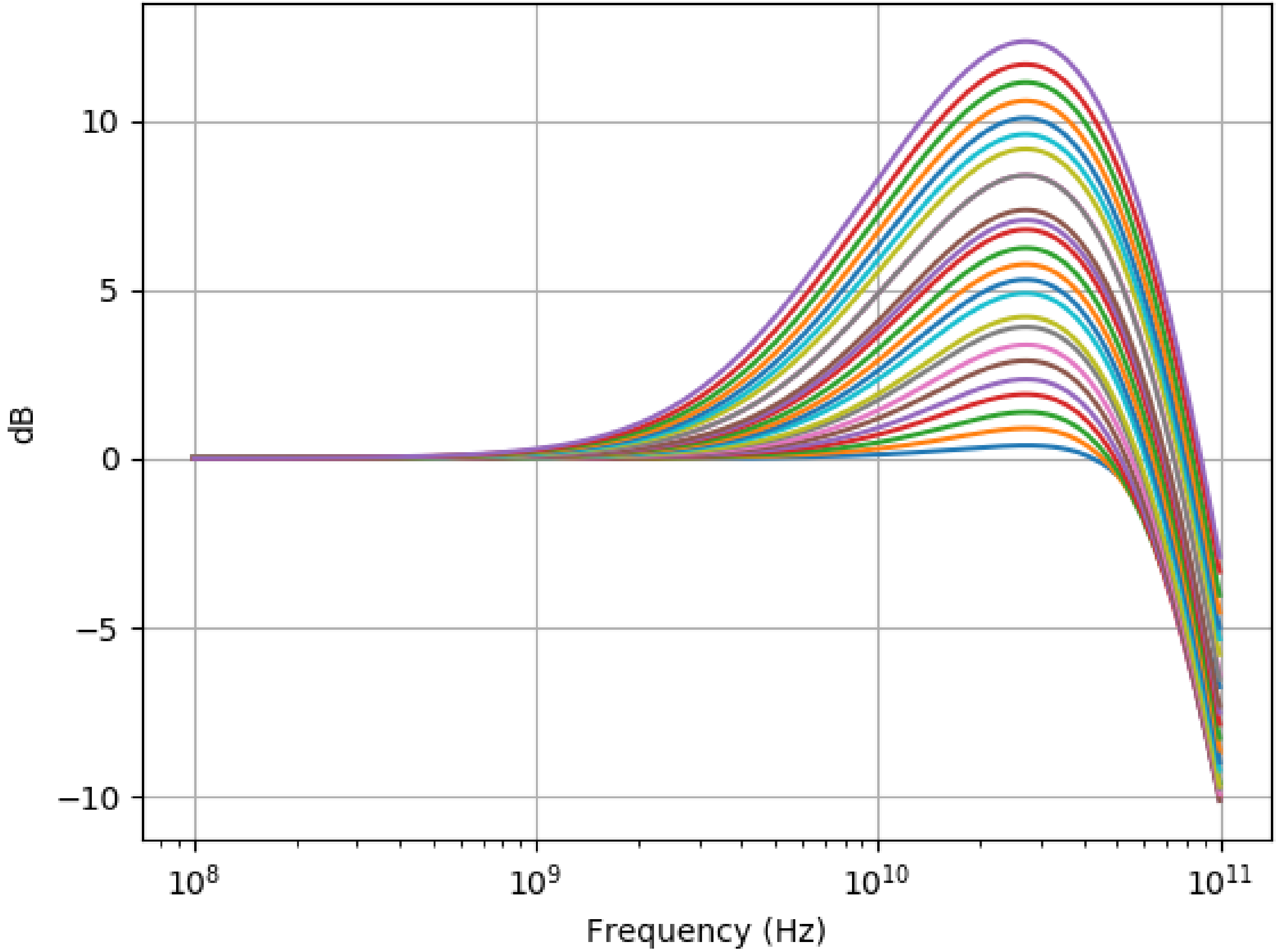


Number of Iteration = 258

$[(f_{pk}, pk), BW]$ = [(27.1e9, 11.7), 50.1e9]

Elapsed time is 13 sec

Examples: EL with 24 pk_{spec} swept



Elapsed time is 493.6 sec

Agenda

- Introduction to Reinforcement Learning
  - Application Examples
- Synthesis of CTLE Modeling
- Spec-Driven w/ Reinforcement Learning
- Examples
- **Summary**

XILINX

Summary

- Reinforcement Learning had been used in many areas and showing its capabilities outperforming human being such as “video games” and “go”.

- Spec-driven reinforcement learning provide a solution to design the architecture spec of CTLE.
  - It is not abnormal to model CTLE with one zero and five poles.
  - It is impractical to do the exhaustive search for the locations of zero and poles to fit the CTLE into specs.

- A few examples have been demonstrated here in this paper to show the efficiency of find the design parameters of CTLE.