



# The Zen Performer®: A New Approach in Deep Coaching

Patrice Robert Nko'o

patnko@stanford.edu, patricenko@gmail.com

Computer Science Department, Stanford University

Stanford  
Science Department

## Overview

Deep Coaching, a Reinforcement Learning Model, called The Zen Performer®, will be used to explore the possibility to reduce the stress generated within an organization because of the lack of leadership.



## Problem

Connecting Human with an RL Model to accelerate the learning process of an agent:

- Modeling RL for the Zen Performer Features
- Modeling a Human Network in the context of parallel coaching
- Connecting two different models

Question: By connecting different models can we gain in term of efficiency (speed learning) ?

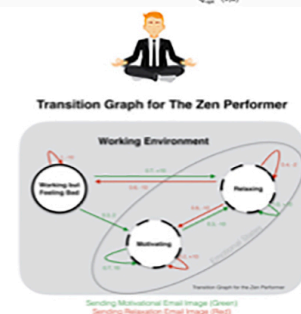
## Framework



## Training

Our agents are trained to maximize their rewards. The Deep Coaching Algorithm used a value iteration from the 1957 bellman equation 1957:

$$V_{opt}^{(i)}(s) \leftarrow \max_{a \in \text{Actions}(s)} \sum_{s'} T(s, a, s') [Reward(s, a, s') + \gamma V_{opt}^{(i-1)}(s')]$$



## Results

RL results without integrating the HCAI Network Data.

Policy Evaluation After 10 Iterations

State	V(s)	Pi(s)
1	101.25	Send Relaxing Image
2	96.56	Send Motivating Image
3	89.80	Send Relaxing Image
4	83.85	Send Motivating Image
5	72.03	Send Relaxing Image
6	64.22	Send Motivating Image
7	44.98	Send Relaxing Image
8	35.58	Send Motivating Image
9	10	Send Relaxing Image
10	0	None
318.06	598.27	53.16%

## Comparative Models

The Zen Performer® Deep Coaching has been compared to Deep Coach, D-COACH, COACH, Advise and CoachAI

Models	Max Score / Rewards	Iterations / Time
Advise	500	300 episodes
COACH	200	10 episodes
D-Coach	500	10 minutes
Zen Performer	598* scale up by 1.5 with HCAI to a reward of 897*	10 iterations

## Future work

Integration of more features within the model to accelerate the speed of learning.

## Discussion

- Which type of AI Coach Platform can improve diversity ?
- How to better assess the societal impact of AI Coaching Platforms ?



## References and Acknowledgement

Thanks to Professor Percy Liang, Dorsa Sadigh, TA Reid Pryzant, and Andrew Zane Tan

Rodrigo Pérez-Dattari, Carlos Celemin, Javier Ruiz-del-Solar and Jens Kober, "Interactive Learning with Corrective Feedback for Policies based on Deep Neural Networks" Dilip Arumugam, Jun Ki Lee, Sophie Saskin, Michael Littman, "Deep Reinforcement Learning from Policy - Dependent Human Feedback" Ahmed Fadhil, Gianluca Schiavo, Yunlong Wang, "CoachAI: A Conversational Agent Assisted Health Coaching Platform" Shane Griffith, Kaushik Subramanian, Jonathan Scholz, Charles Isbell and Andrea Thomaz, "Policy Shaping: Integrating Human Feedback with Reinforcement Learning"