Exploration and Fairness in Infinite Armed Bandit Problems

Akshay Channesh and **Venkateshwar Ragavan**
University of Illinois Chicago
achann4,vragv2@uic.edu

Abstract

1 Introduction

(Parikh, 2019) (Parikh, 2010) (Blum et al., 2020)

2 Related Work and Preliminary Concepts

2.1 Infinite Multi-Armed-Bandit

2.2 Regret Bound

2.3 UCB

2.4 Thompson Sampling

2.5 Fairness

3 Proposed Method

3.1 Curiosity

3.2 Surplus

4 Experiments and Discussion

4.1 The Agents

1. Random Agent: This agent picks an unseen arm half the times and a random seen arm otherwise.

2. Always New Arm Agent: This agent always picks an unseen arm.

3. Naive UCB Agent: This agent picks a new unseen arm with $p = 1\%$ and every other time it picks among the known arms using UCB.

4. Naive Thompson Sampling Agent: This agent picks a new unseen arm with $p = 1\%$ and every other time it picks among the known arms using Thompson Sampling.

5. Fair UCB Agent: This agent picks a new unseen arm with $p = 1\%$ and every other time it picks among the known arms using UCB while also ensuring α-Fairness among known arms.

6. Fair TS Agent: This agent picks a new unseen arm with $p = 1\%$ and every other time it picks among the known arms using Thompson Sampling while also ensuring α-Fairness among known arms.

7. Surplus Curiosity UCB Agent: This agent picks a new unseen arm with a probability based on the surplus, and every other time it picks among known arms using UCB.

8. Surplus Curiosity TS Agent: This agent picks a new unseen arm with a probability based on the surplus, and every other time it picks among known arms using Thompson Sampling.

4.2 Environments

1. Uniform Environment

2. Increasingly Better Environment

3. Progressively Worse Environment

4.3 Results

4.3.1 Uniform Environment

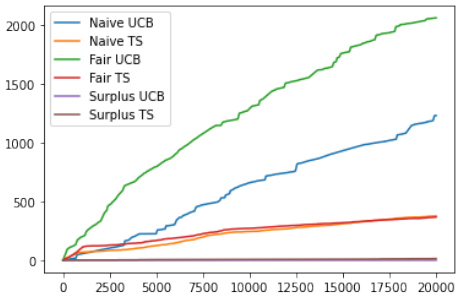

Figure 1: Uniform Environment. Linear-Linear Graph

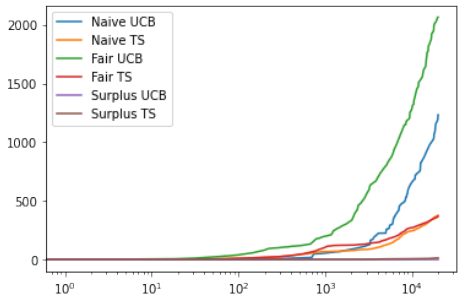

Figure 2: Uniform Environment. Log-Linear Graph

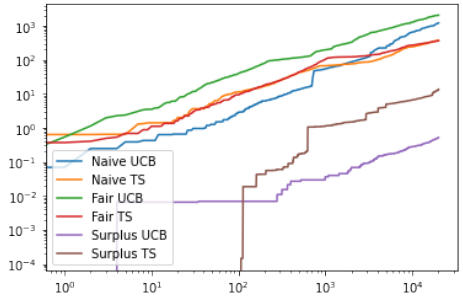

Figure 3: Uniform Environment. Log-Log Graph

4.3.2 Increasingly Better Environment

4.3.3 Progressively Worse Environment

4.4 Discussion

5 Conclusion

References

Avrim Blum, John Hopcroft, and Ravindran Kannan. 2020. *Foundations of Data Science*. Cambridge University Press, Cambridge.

Prashant Parikh. 2010. *Language and Equilibrium*. MIT Press.

Prashant Parikh. 2019. *Communication and content*. Number 4 in Topics at the Grammar-Discourse Interface. Language Science Press, Berlin.

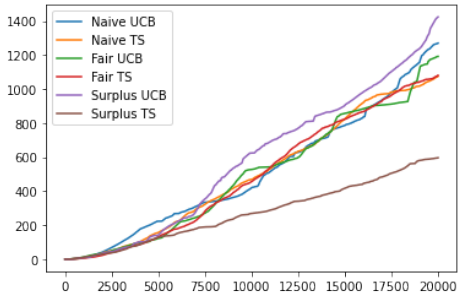

Figure 4: Increasingly Better Environment. Linear-Linear Graph

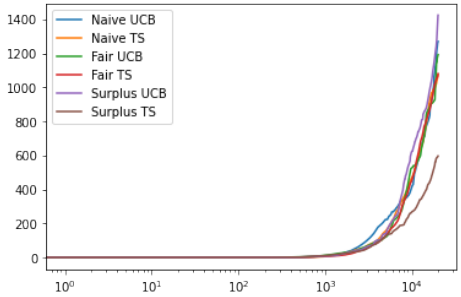

Figure 5: Increasingly Better Environment. Log-Linear Graph

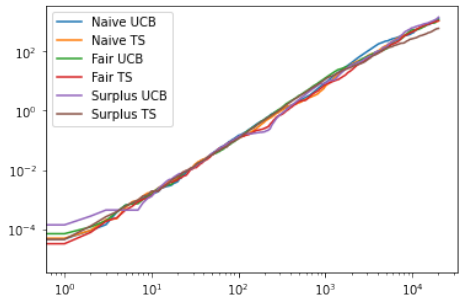

Figure 6: Increasingly Better Environment. Log-Log Graph

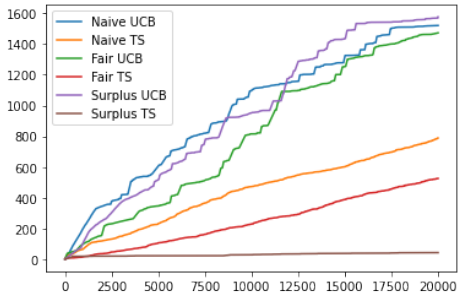

Figure 7: Progressively Worse Environment. Linear-Linear Graph

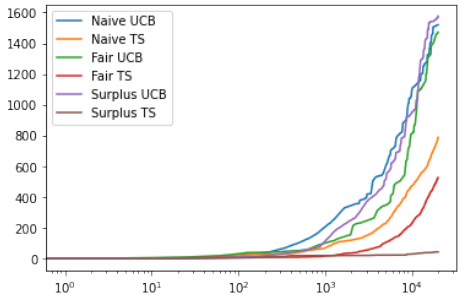

Figure 8: Progressively Worse Environment. Log-Linear Graph

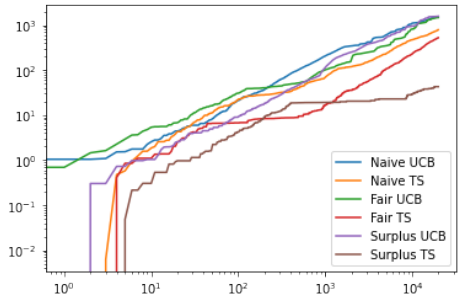

Figure 9: Progressively Worse Environment. Log-Log Graph