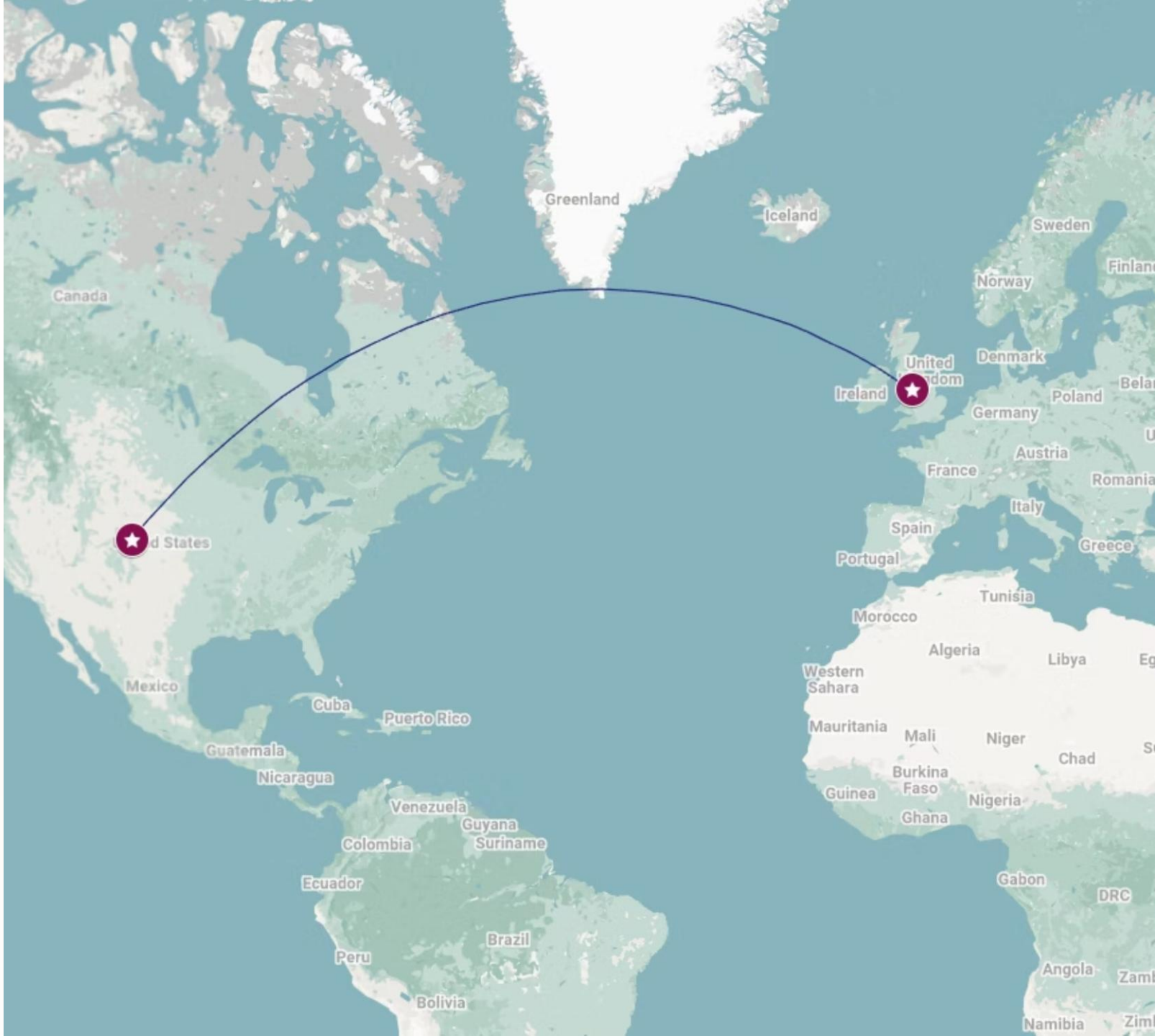
Social Welfare Under Heterogeneous Time Preferences

Sarvin Bahmani

University of Liverpool

Soumyajit Paul, Sven Schewe,
Shadi Kalat, Ashutosh Trivedi.



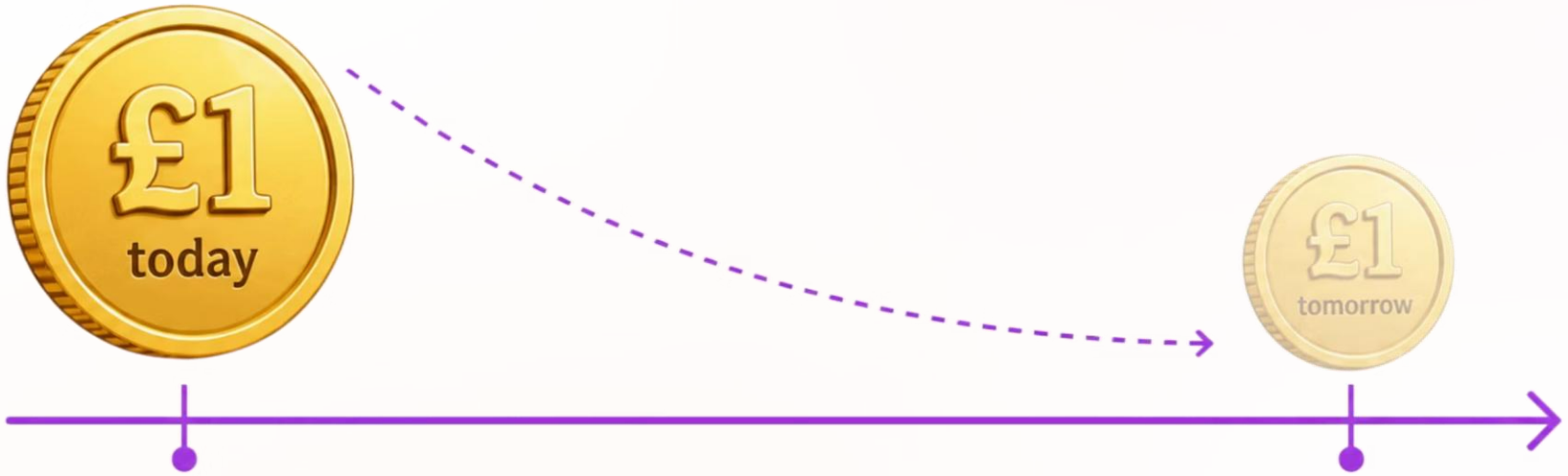
Why Heterogeneous Time Preferences?

Climate policy

Infrastructure

Investment

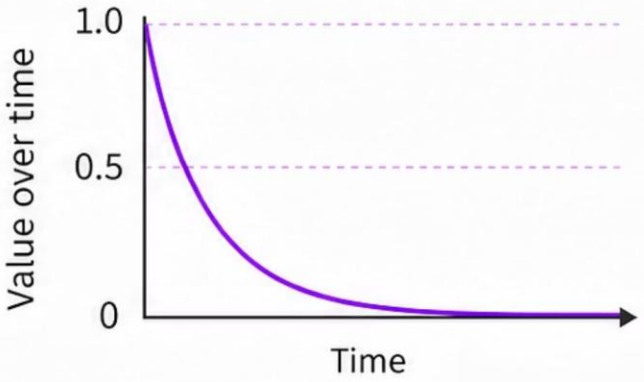
The Value of Delayed Rewards



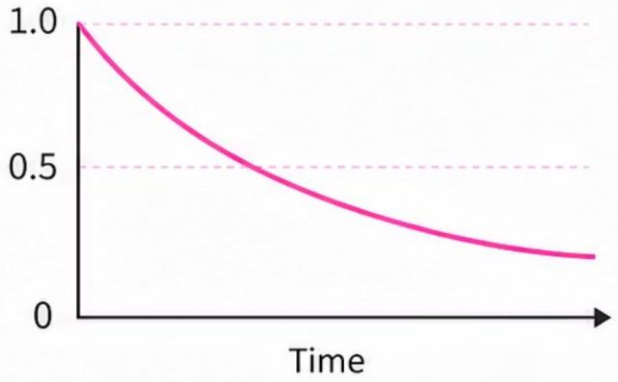
A reward becomes **less valuable** when it is delayed.

Do People Value the Future Differently?

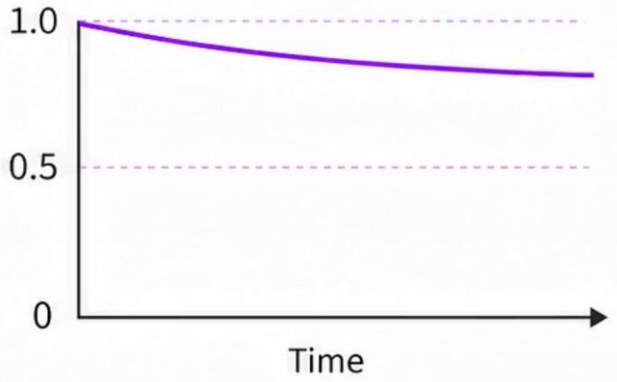
Yes! People value the future differently.



Low discounting
(impatient)



Moderate discounting



High discounting
(patient)

One Decision-Maker

Agent , Decision maker

Multiple Principals

Principals, stakeholders

Social Welfare Objective

Principal i's payoff under policy π

$$V_i^{\pi} = \mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty} \lambda_i^t \, R_i^t\right]$$

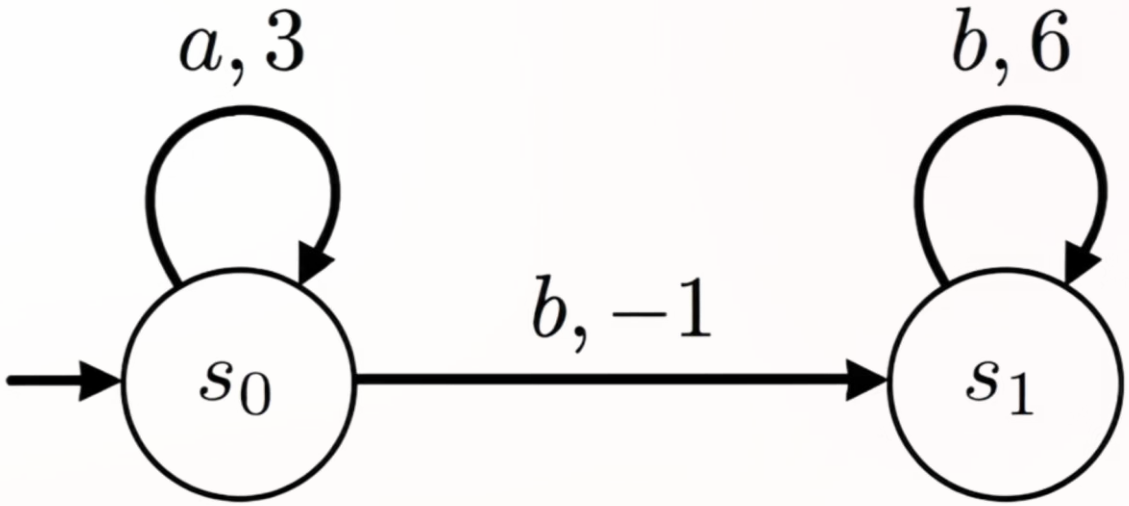
Social Welfare

$$\mathrm{SW}(\pi) = \sum_i V_i^{\pi}$$

Optimisation Objective

$$\max_{\pi} \mathrm{SW}(\pi)$$

Alice and Bob, Open a Restaurant!



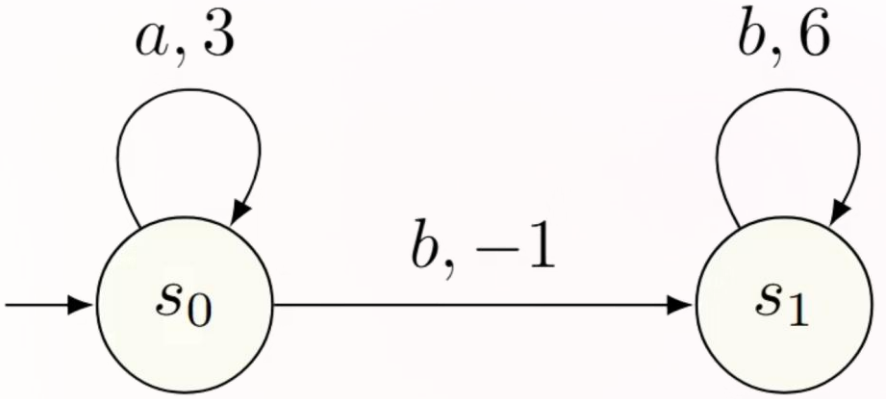
Patient investor

Alice $\boldsymbol{\lambda}$ **: 2/3**

Impatient investor

Bob $\boldsymbol{\lambda}$ **: 1/3**

Social Welfare in different strategies



Policy 1: a^ω

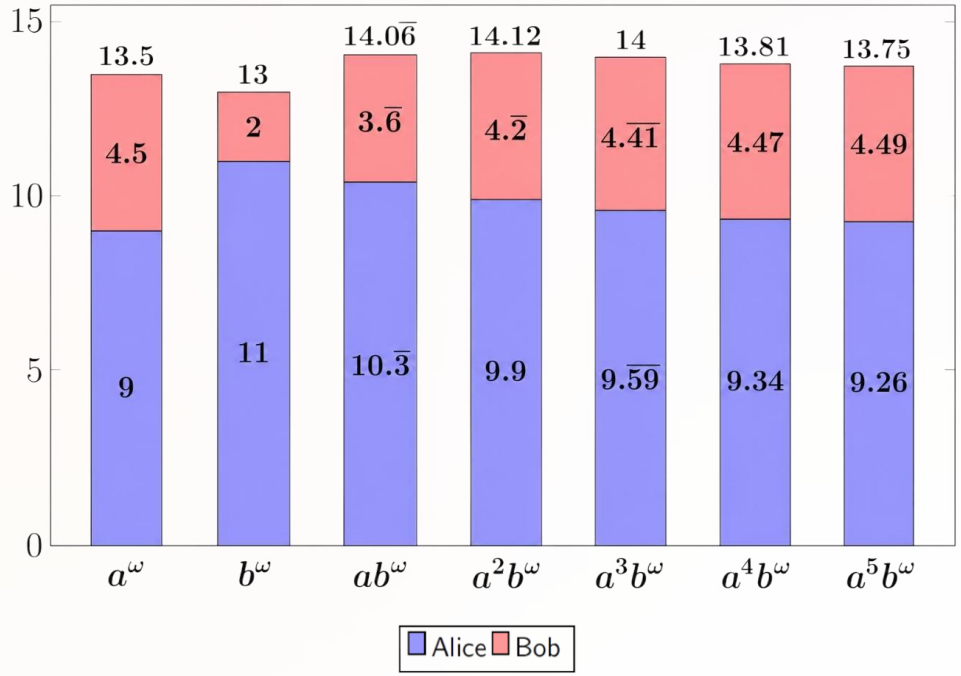
individual payoff: $\frac{3}{1-\lambda}$

Policy 2: b^ω

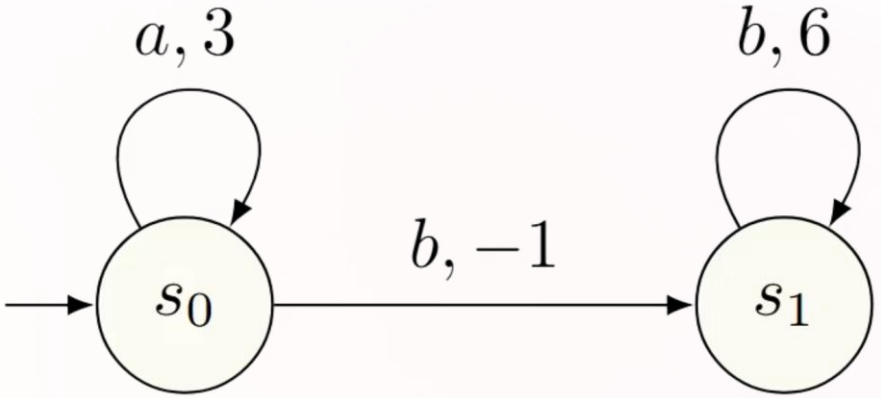
individual payoff: $-1 + 6\frac{\lambda}{1-\lambda}$

Policy 3: $a^k\ b^\omega$

individual payoff: $3\frac{1-\lambda^k}{1-\lambda} - \lambda^k + 6\frac{\lambda^{k+1}}{1-\lambda}$



Social Welfare in different strategies



Policy 1: a^ω

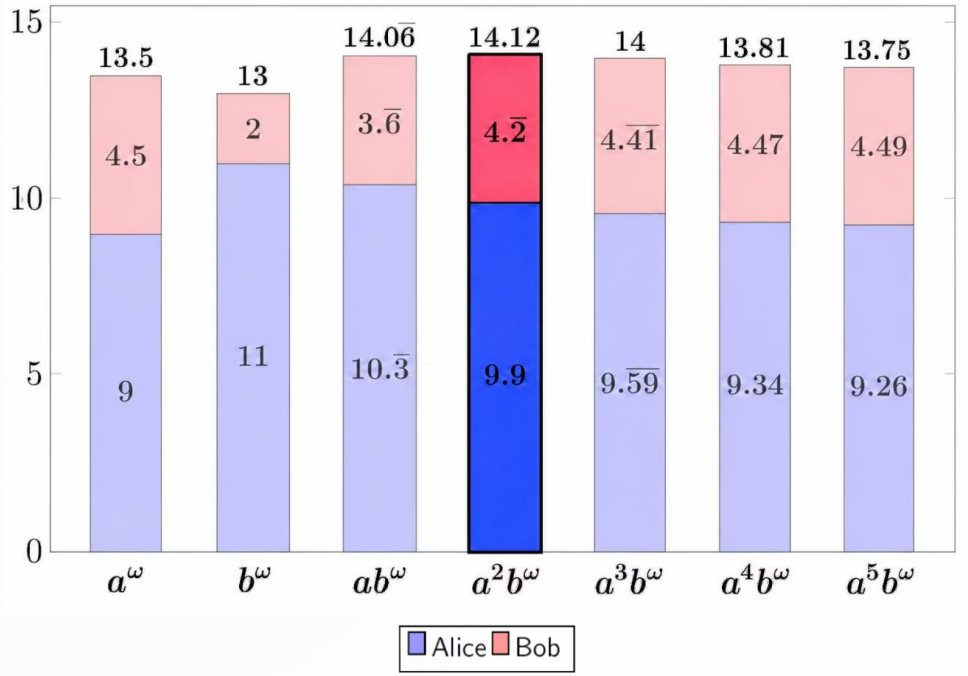
individual payoff: $\frac{3}{1-\lambda}$

Policy 2: b^ω

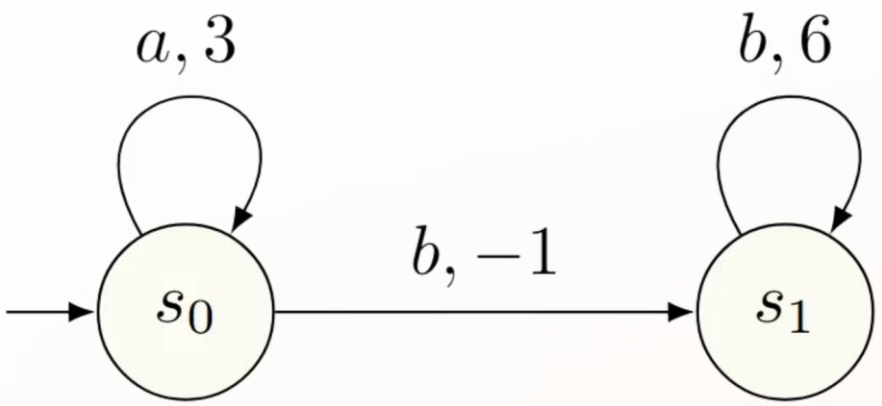
individual payoff: $-1 + 6\frac{\lambda}{1-\lambda}$

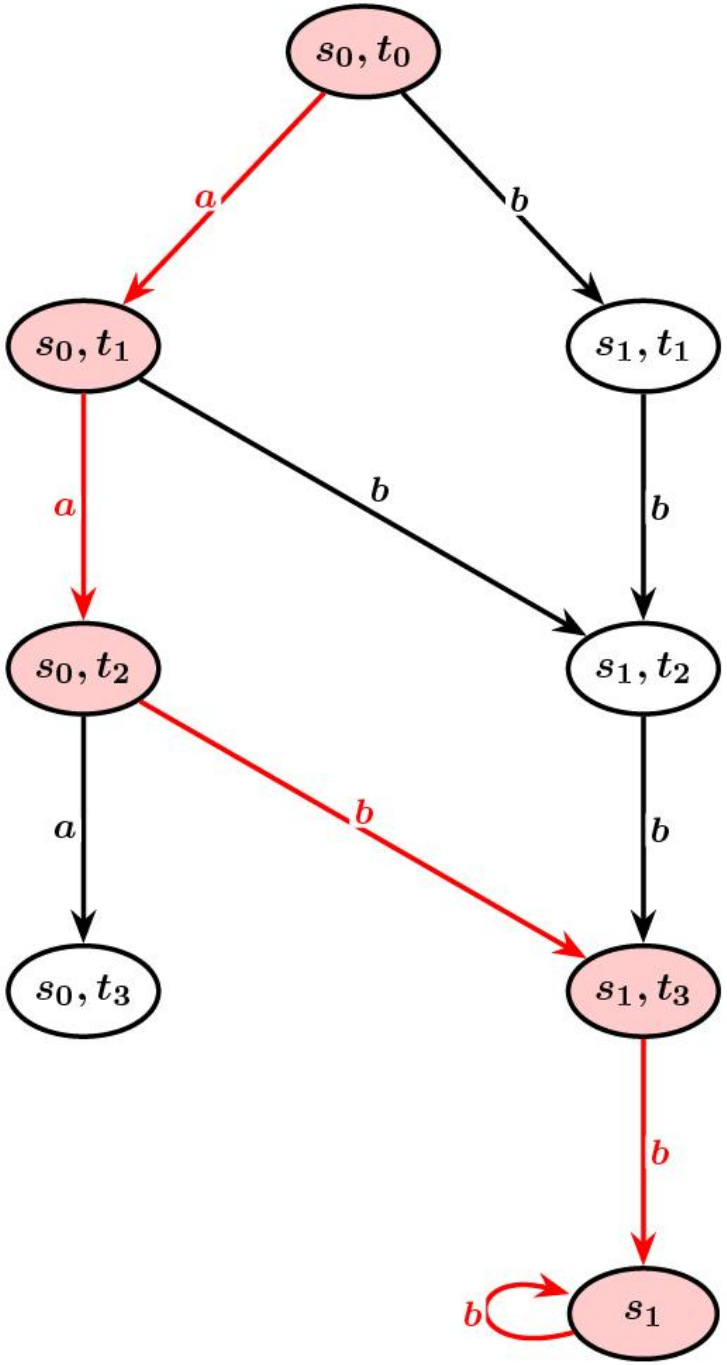
Policy 3: $a^k\ b^\omega$

individual payoff: $3\frac{1-\lambda^k}{1-\lambda} - \lambda^k + 6\frac{\lambda^{k+1}}{1-\lambda}$

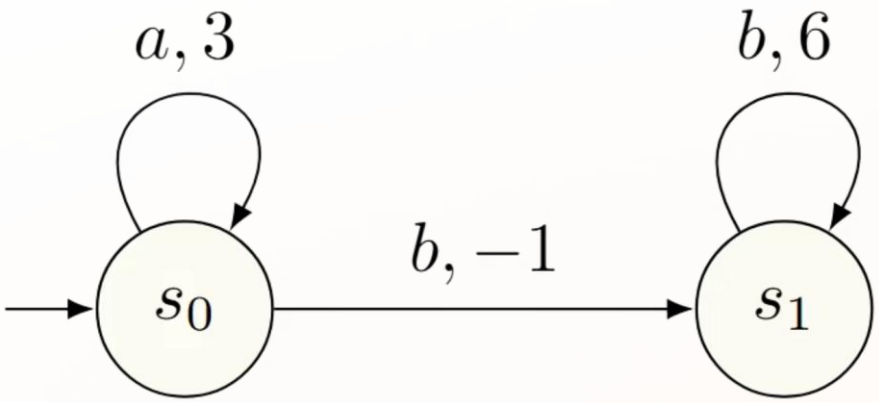


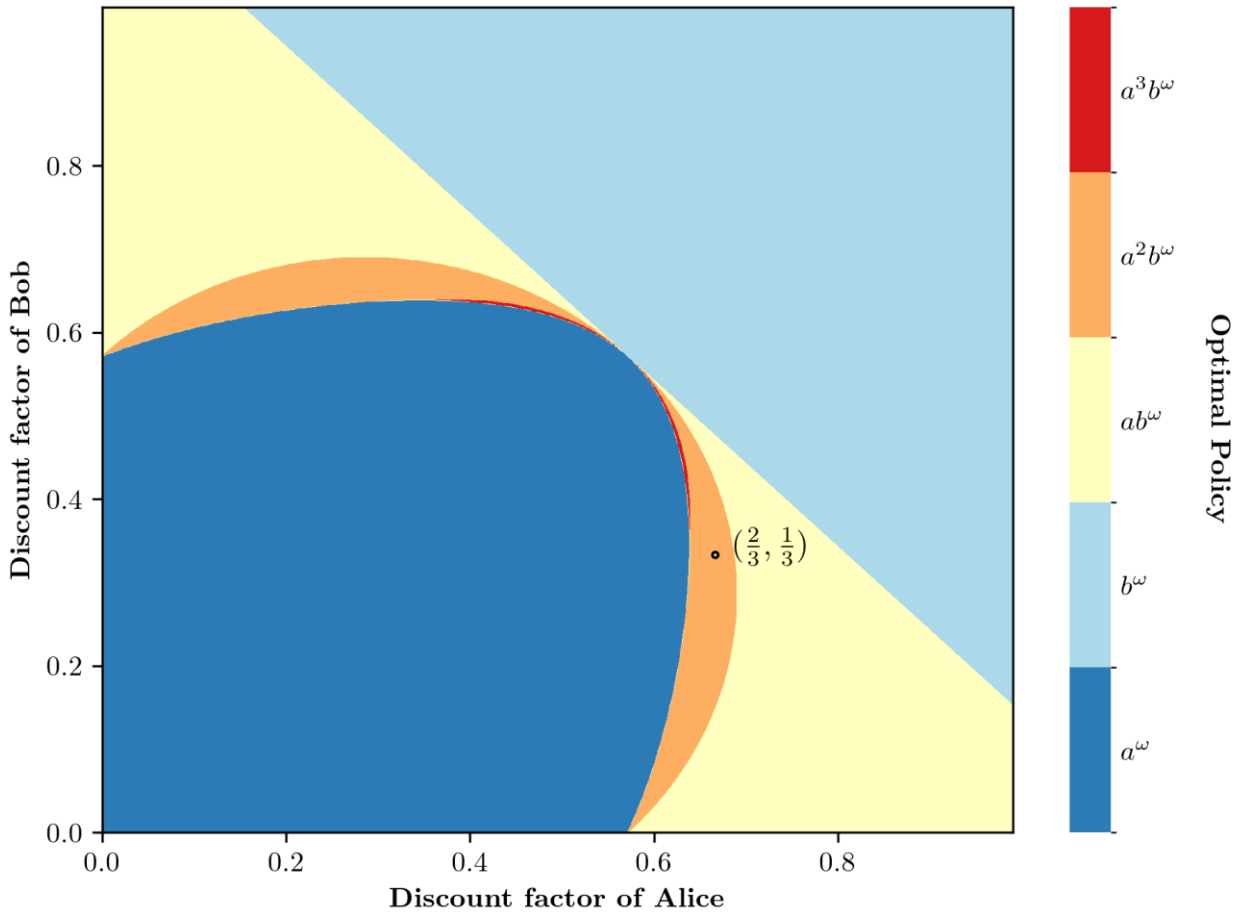
Social Welfare
in different strategies





Heterogeneous discounting turns time itself into **Strategically relevant memory.**





How to get the best Social Welfare?

Bad News

- Pure stationary strategies?
- Randomised strategies?

Good News

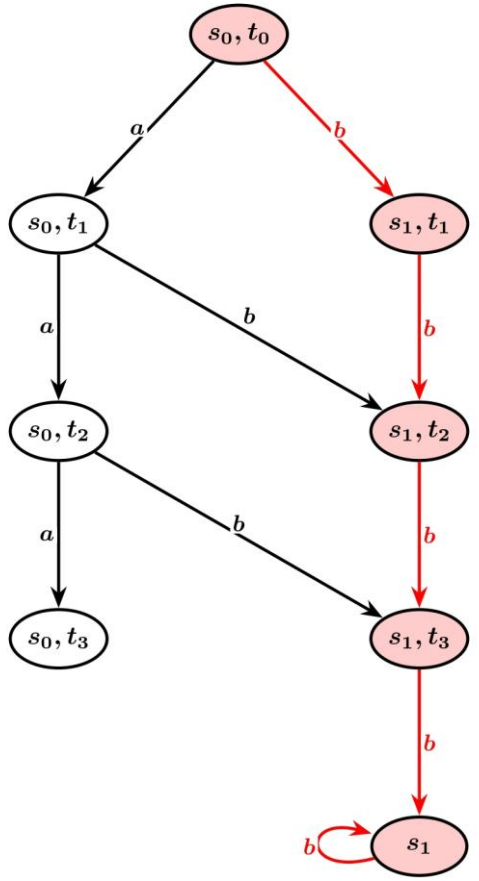
- ✓ Optimal strategies exist in a simple class.
- ✓ Pure finite-memory counting strategies

Simple stationary behavior is Hard

Structured finite memory is easy.

Long-run Behavior: Lexicographic Tail

Order principals by patience: $\lambda_0 > \lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_n$



1. **Optimize the MDP for the most patient principal.**
   a. **Restrict to actions that are optimum.**
2. **Break ties using the next most patient principal.**
3. **Continue until last principals.**

M_0 original MDP → M_1 optimal for P_0 → M_2 tie-break P_1 → . . . → M_n long-run tail

Deviation from Long Term Strategy

M_n
long-run tail

For Principal$_i$:

$$\Delta_0(s,a,i) = R(s,a,i) + \lambda_i \sum_{s'} p(s' \mid s,a) V_i(s') - V_i(s)$$

After t steps

$$\Delta_t(s,a,i) = \lambda_i^t \Delta_0(s,a,i)$$

Social effect

$$\sum_i \Delta_t(s,a,i) = \sum_i \lambda_i^t \, \Delta_0(s,a,i)$$

After a finite horizon, deviations cannot help

M_n
long-run tail

for every deviation, there exists a finite horizon k such that, for all $t \geq k$:

$$\Sigma_i \, \Delta_t(s, a, i) \geq 0$$

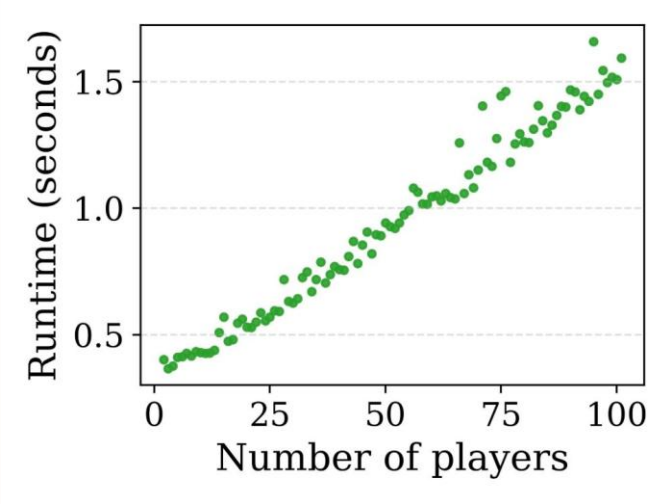
Why κ is small under spacing?

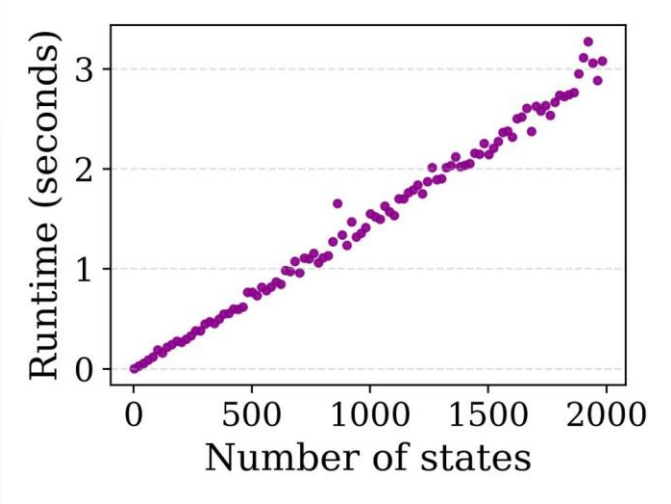
A typical bound has the shape $k = O\left(\frac{\log(\text{value ratio})}{\log(\lambda_i/\lambda_{i+1})}\right)$

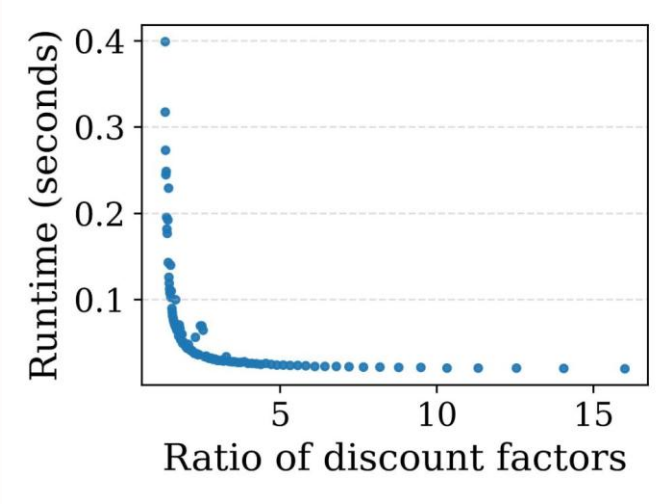
Outcome

Only the prefix before K needs time-dependent optimization; after that, follow the stationary tail.

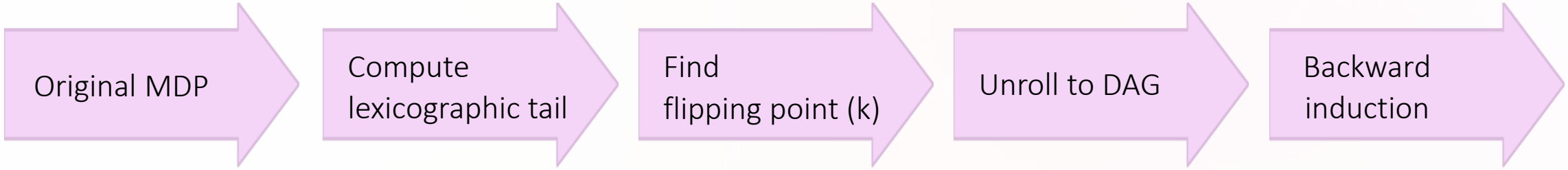
Scalability Trends







Synthesis Algorithm



Deviation rewards: $R_{\mathrm{k}}^{\Delta}(s,a) = \sum_i \lambda_i^{\mathrm{k}}\, \Delta_0(s,a,i)$

$$\pi^* = \underbrace{[\text{finite prefix}]}_{\text{counting memory}} \cdot \underbrace{[\pi_{\text{long}}^{\omega}]}_{\text{Stationary tail}}$$

Thank you!

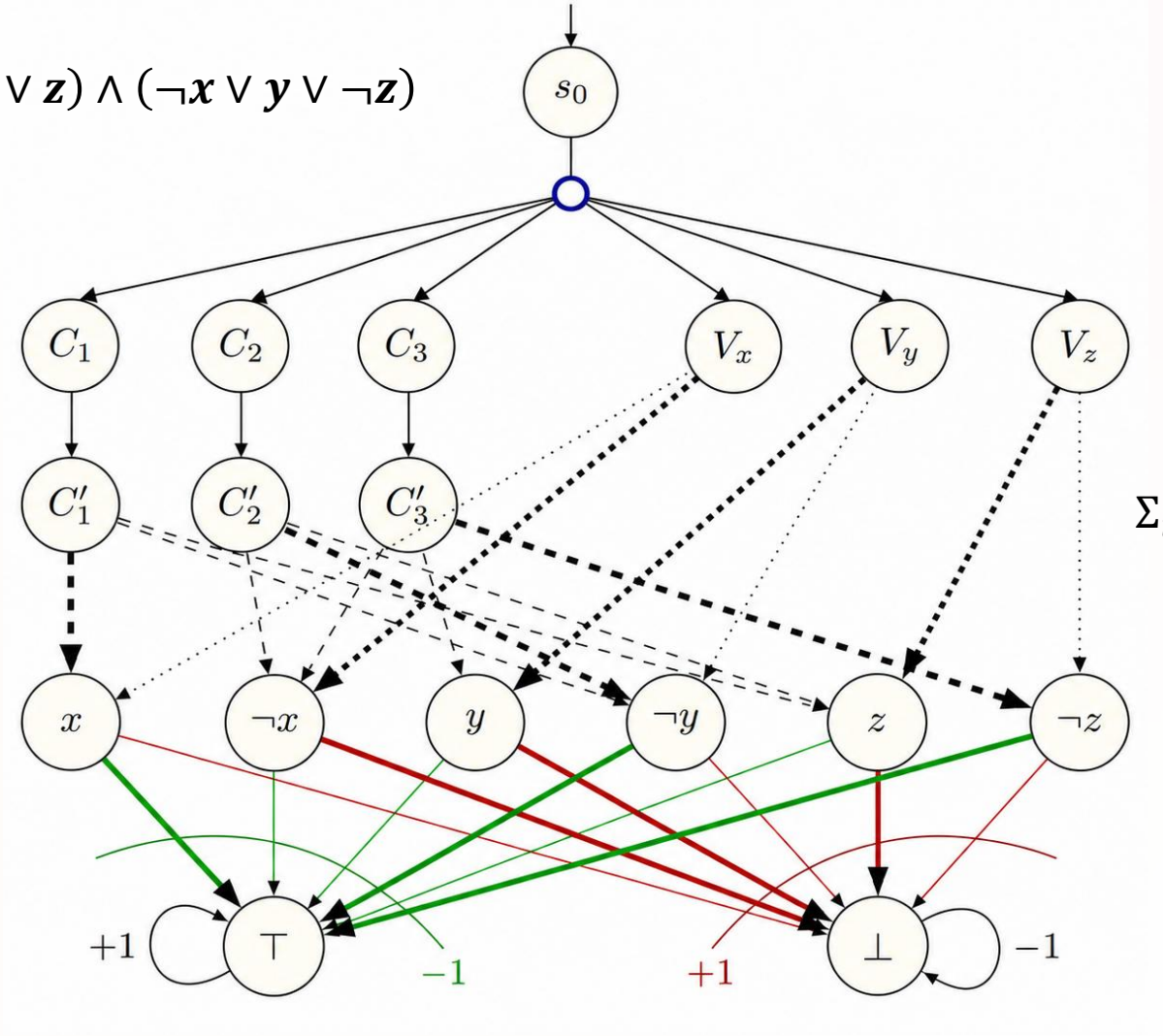
NP-Hardness

$\boldsymbol{\phi} = (\boldsymbol{x} \vee \neg \boldsymbol{y} \vee \boldsymbol{z}) \wedge (\neg \boldsymbol{x} \vee \neg \boldsymbol{y} \vee \boldsymbol{z}) \wedge (\neg \boldsymbol{x} \vee \boldsymbol{y} \vee \neg \boldsymbol{z})$

$\lambda_0 = 0.54$

$\lambda_1 = 0.40$

$$\Sigma_i \left(0 + 0 \cdot \lambda_i + 0 \cdot \lambda_i^2 - {\lambda_i}^3 + \frac{\lambda_i^4}{1 - \lambda_i} \right)$$



$$\Sigma_i \left(0 + 0 \cdot \lambda_i + {\lambda_i}^2 - \frac{\lambda_i^3}{1 - \lambda_i} \right)$$