Prefetching;
A Markov Decision Process Model

Kausthub Keshava

MS16010

A dissertation submitted for the partial fulfilment of BS-MS dual degree in Science

Indian Institute of Science Education and Research, Mohali

April 2021

Certificate of Examination

This is to certify that the dissertation titled "**Prefetching; A Markov Decision Process Model**" submitted by Mr. Kausthub Keshava (Reg. No. MS16010) for the partial fulfilment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr.Abhik Ganguli	Dr. Pranab Sardar	Dr. Varadharaj R. Srinivasan (Supervisor)

Dated: April 28, 2021

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Prof. Alain Jean-Marie and Prof. Sara Alouf at the National Institute for Research in Digital Science and Technology (Inria), France, and Dr. Varadharaj R. Srinivasan at the Indian Institute of Science Education and Research, Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of work done by me and all sources listed within have been detailed in the bibliography.

Kausthub Keshava
(MS16010)

Dated: April 28, 2021

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Varadharaj R. Srinivasan
(Supervisor)

Acknowledgements

I thank Prof. Alain Jean-Marie and Prof. Sara Alouf for their guidance and support without which this thesis would not be possible. I also thank Dr. Varadharaj R. Srinivasan for his encouragement and supervision of my progress as my local guide. I would like to thank the Network Engineering and Operations (NEO) team at the National Institute for Research in Digital Science and Technology (Inria), France, for providing me the opportunity to do this project. I also thank Inria for the financial support that they provided during the course of my collaboration with NEO. I am grateful to the authors of the texts that I have referred to, for having written them beautifully, and showing me the way into the wonderful world of mathematics and optimization.

I am thankful to the faculty in the mathematics department for all they have taught me through my years in IISER, especially Dr. Neeraja Sahasrabudhe who furthered my interest in random processes. I am eternally grateful to IISER Mohali for providing me with an environment to learn and grow like none other. I am thankful for the Library at IISER Mohali for providing an amazing collection of books and a conducive environment for thesis work. I also thank my dear friends Abhijit, Nandan, Nikhil, Puneeth, Ruchira, and Sasank for their love and emotional support during this entire project. Finally, I would like to thank my family Dr. B S Keshava, Dr. Nandini Keshava, Mrs. Namratha Keshava, and Mr. Adishesh Kishore for always being there whenever I needed them.

List of Figures

List of Tables

List of notation

Table 1: Table of Symbols

No.	Symbol	Meaning
1	$\mathcal{T}_{p,d}$	Set of rooted, ordered, and marked trees of depth d with nodes having between 1 and p sons
2	$\mathcal{L}_k$	$= \mathcal{T}_{p,k} \cup (\mathcal{T}_{p,k} \times \mathcal{T}_{p,k}) \cup \ldots \cup (\mathcal{T}_{p,k})^p$
3	(μ, s)	Representation of a tree t in $\mathcal{T}_{p,d}$
4	μ	Mark of the root of tree and takes values in $\{0, 1\}$
5	s	List of elements in $\mathcal{L}_{d-1}$
6	n	Time step of process
7	t_a	Tree after controller makes an action
8	t_b	Subtree of t_a that the surfer moves to
9	U	Usable state space of the MDP
10	$L(t)$	Set of leaves of tree t
11	$\mathcal{D}(t)$	Set of possible discoveries of tree t
12	maxnode(d)	The maximum number of nodes of tree in $\mathcal{T}_{p,d}$
13	k	Budget given to controller
14	$X_d(m)$	Number of trees in $\mathcal{T}_{p,d}$ with m nodes
15	N	Time horizon of the MDP
16	ρ	Stationary probability matrix of appropriate size

Table 2: Table of Symbols Ctd.

No.	Symbol	Meaning
17	$\mathcal{SD}(t)$	Set of subtree discoveries of tree t
18	x^+	x for $x > 0$, and 0 otherwise.
19	$\mathbb{H}_p$	Sum of harmonic series until p, i.e., $\sum\limits_{f=1}^{p} f^{-1}$
20	minnode(d)	The minimum number of nodes of tree in $\mathcal{T}_{p,d}$ which is $d+1$
21	$d_1(t)$	Set of depth 1 nodes of tree t
22	$\mathbb{H}_{pk}$	$\sum_{e=1}^{p}(e-k)^+/e$
23	T_{kbig}	Set of trees with number of unmarked sons greater than or equal to the budget k
24	T_{ksmall}	Set of trees with number of unmarked sons less than the budget k
25	$\delta(l,t)$	Depth of the node l in tree t
26	$sib(l)$	Number of unmarked siblings of node l which takes values from $0, 1, \cdots, p-1$
27	$tsib(l)$	Total number of siblings of node l, which takes values from $0, 1, \cdots, p-1$

Contents

Abstract

Prefetching is a technique used to boost computer execution performance by fetching instructions or data before it is actually needed. Constraints on network bandwidth and memory lead to a choice being made in terms of the specific data to be prefetched. Markov Decision Processes are a valuable tool to model stochastic decision making. One can view the prefetching process as a random process of a surfer moving on a graph and a controller trying to ensure that the surfer lands on a prefetched vertex. This is analogous to the well-known pursuit-evasion game in graph theory. Our aim is to find the optimal policy for the controller and explore the characteristics of such a policy. Throughout this thesis, we analyse and study the properties of the prefetching process modelled on a tree (rooted acyclic graph) as an MDP. Using the Bellman Optimality equations, we solve for the optimal policy of the prefetching MDP for different criteria. In the finite horizon criterion, we obtained a granular greedy optimal policy. We implemented policy iteration for the average costs infinite horizon criterion and converged to the optimal policy of the finite horizon MDP. The optimal policy is dependent on the specific shape and size of the trees. We structured the state space in a manner that eased the search for the optimal value function and corresponding optimal policy given a particular state.

Chapter 1

Introduction

This thesis is an attempt at characterising and analysing optimal policies for the prefetching process with uncertainty. We use the framework of a Markov Decision Process to carry out our analysis.

First, we shall establish some intuition for the concept of an MDP and prefetching. Following this, we will go through the definition and objects of an MDP. This thesis focuses on two specific criteria of MDPs and hence a brief outline of the theory of these particular criteria have been included.

After establishing existential theorems and methodologies to compute an optimal policy, we shall understand how the prefetching process can be viewed as an MDP. We construct the prefetching process on a rooted tree of certain depth with every node having a maximum number of sons.

Next we shall characterise the optimal policy in the finite horizon expected costs criterion for depth 1 trees. We use various proof techniques to establish that a proposed policy is optimal. We then move on to depth 2 trees, beginning with budget 1 and then to budget 2. The results of the finite horizon case provide insights into possible policies to explore in the infinite horizon average costs criterion. Chapter 6 extensively deals with the Markov chain of tree shapes and then tackles specific cases of the infinite horizon MDP.

It is assumed that the reader has a basic understanding of Markov chains, limits, and combinatorics.

1.1 Prefetching

Prefetching is the process of loading of a resource before it is required to decrease the time spent in waiting for that resource. An example of prefetching is - instruction prefetching where a CPU caches data and instruction blocks before they are executed. Prefetching functions often make use of a cache to store the "prefetched" resources. Web browsers employ prefetching by preloading commonly accessed pages. When the user navigates to the page that is prefetched, it loads quickly because the browser is pulling it from the cache, rather than from a distant sever. Some browser plugins download all of the pages that have been hyperlinked in an attempt to speed up the browser. However this comes with an increase in bandwidth usage. The problem to solve would be : given a constraint on bandwidth, what are the specific resources to prefetch to minimize the possibility of a user accessing a resource that is not prefetched into the cache. Figure 1.1 provides a visual depiction of the prefetching process.

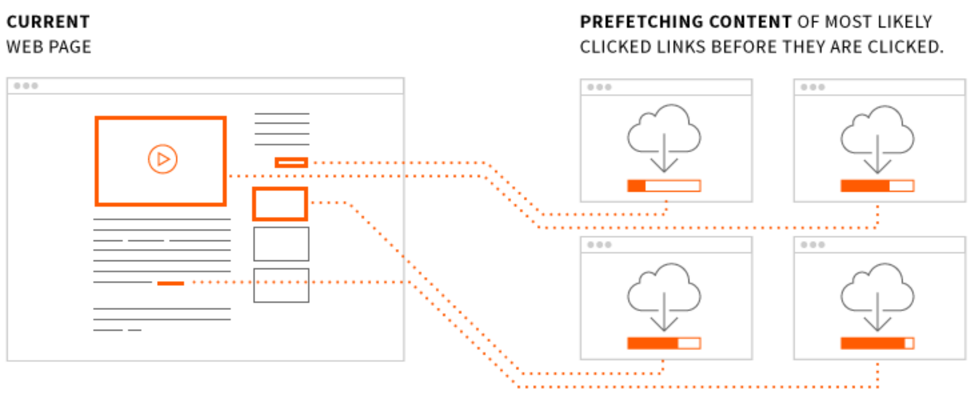

Figure 1.1: Prefetching

Source: maxcdn.com/what-is-prefetching

1.2 Markov Decision Process

A Markov decision process is a discrete time stochastic control process. It is a framework that can model decision making in situations where there is randomness in outcomes. MDPs are an extension of Markov chains with the addition of actions (the "decision") and rewards. At each time step, the process is in some state, and the decision maker may choose any action that is available in that state. Then the process moves into a new state and gives the decision maker a corresponding reward (or cost) for his action. The probability of the process transitioning into a new state depends on the initial state and the action taken at that initial state. We obtain a simple Markov chain from an MDP if there is only one action available for all states and a constant reward of zero for all states and actions. An example of an MDP with three states and two actions is depicted in Figure 1.2. The possible rewards are $\{+1, +2, -10\}$ depending on the state and action taken.

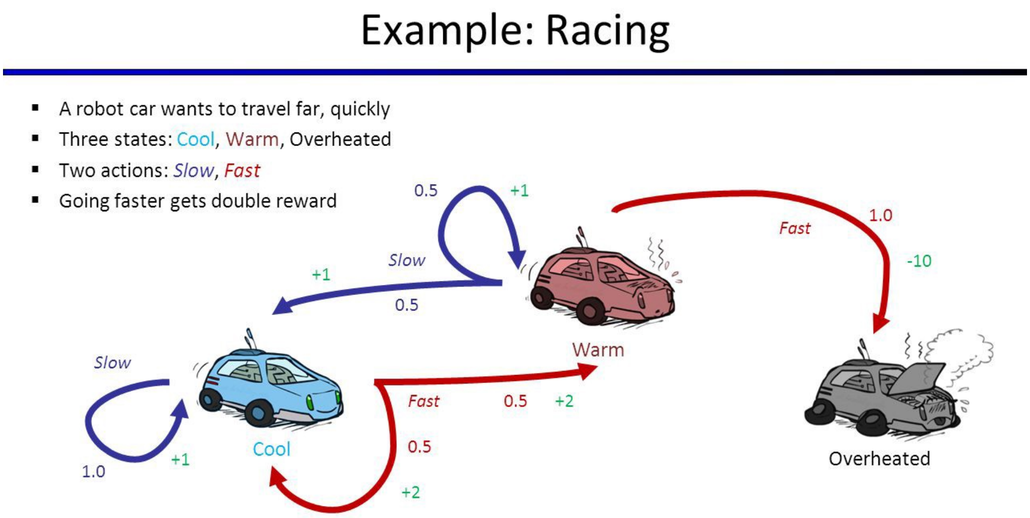

Figure 1.2: Markov Decision Process

Source: medium.com/MDP

Chapter 2

Preliminaries : MDP Theory

2.1 Finite Horizon Expected Reward Criterion

The theory included in this chapter follows the notation and logic from Chapters 2 and 4 of the book titled *Markov Decision Processes* by Martin L Puterman [1] and Chapter 5 of the book titled *Introduction to Stochastic Dynamic Programming* by Sheldon Ross [2]. First, let us formally define a Markov Decision Process, a decision rule, a policy, value function of a state, and expected total reward at a particular time.

Definition 2.1.1 (Markov Decision Process)**.** A collection of objects $\{T, S, A_s, p_t(.|s,a), r_t(s,a)\}$ is an MDP where:

- T = Set of decision epochs/Time steps. $T = \{1, \cdots, N\}, N < \infty$ or $T = [1, \infty)$.
- S, a finite, countable, or compact subset of $\mathbb{R}^n$, is a set of all possible states. Let the random variable s_t denote the state of the process at t. $s_t \in S$ for all $t \in T$.
- A_s, a finite, countable, or compact subset of $\mathbb{R}^n$, is the set of all allowable actions in state s. Let $A = \cup_{s \in S} A_s$. Let the random variable a_t denote the

action taken at t. $a_t \in A$ for all $t \in T$.

- $r_t(s,a) : T \times S \times A \to \mathbb{R}$ is the reward received at decision epoch t as a result of choosing action $a \in A_s$ while in state $s \in S$.

- $p_t(.|s,a) : T \times S \times S \times A \to [0,1]$ is the transition probability that determines the next state. $P(s_{t+1} = s'|s_t = s, a_t = a) = p_t(s'|s,a)$.

Definition 2.1.2 (Decision rule)**.** At time $t \in T$, a decision rule d_t prescribes a rule for action selection in each state. There are four types of rules as follows:

- (MD) Markovian Deterministic : $d_t : S \to A$, $d_t(s) \in A_s$ denotes the action chosen at time t in state s.

- (HD) History Dependent Deterministic : $d_t : H_t \to A$. $h_t = (s_1, a_1, \cdots, s_t)$, and H_t denotes the set of all histories h_t. $d_t(h_t) \in A_{s_t}$.

- (MR) Markovian Randomized : $d_t : S \to \mathscr{P}(A)$. Here $\mathscr{P}(A)$ is the set of all probability distributions on A. $d_t(s_t) = q_{d_t(s_t)}(.) \in \mathscr{P}(A_{s_t})$.

- (HR) History Dependent Randomized : $d_t : H_t \to \mathscr{P}(A)$. $q_{d_t(h_t)}(.) \in \mathscr{P}(A_{s_t})$.

Definition 2.1.3 (Policy)**.** A policy is a sequence of decision rules for each decision epoch. $\pi = (d_1, d_2, \cdots, d_{N-1})$.
Let $\prod^K$ be the set of all policies of class $K \in \{HR, MR, HD, MD\}$. For instance,

- $\pi \in \prod^{HR}$ means d_t is a Randomized history dependent decision rule for all $t \in T$.

Definition 2.1.4 (Value function of a state)**.** The value function of a state s for a policy $\pi \in \prod^{HR}$ is defined as,

$$v_N^\pi(s) = E_s^\pi \left[\sum_{t=1}^{N-1} r_t(s_t, a_t) + r_N(s_N) \right]. \tag{2.1}$$

where $r_N(s_N)$ (Scrap Value) is the reward received when the process ends at N in some state. The expectation is conditioned on $s_0 = s$ and the process evolves according to the policy π.

We seek a policy π^* for which,

$$v_N^{\pi^*}(s) \geq v_N^{\pi}(s) \text{ for all } s \in S \text{ and for all } \pi \in \prod^{HR}. \tag{2.2}$$

Definition 2.1.5 (Expected total reward at time t)**.** The expected total reward at time t of a history h_t is given by $u_t^{\pi}(h_t)$ where

$$\begin{aligned} u_t^{\pi}(h_t) &= E_{h_t}^{\pi}\left[\sum_{n=t}^{N-1} r_n(X_n, Y_n) + r_N(X_N)\right] \\ u_N^{\pi}(h_N) &= r_N(s) \text{ when } h_N = (h_{N-1}, a_{N-1}, s). \end{aligned} \tag{2.3}$$

Here the expectation is conditioned on the history h_t and the process evolves according to policy π.

Note the equivalence between the expected total reward at time t and the value function which is just $u_t^{\pi}(h_t)$ for $t = 1$. Thus we can use these terms interchangeably, unless specified otherwise. We now state a lemma that is the essence of dynamic programming.

Lemma 2.1. *(Multistage to Single stage) For $\pi \in \prod^{HD}$, we have the following simplification for $u_t^{\pi}(h_t)$*

$$u_t^{\pi}(h_t) = r_t(s_t, d_t(h_t)) + \sum_{j \in S} p(j|s_t, d_t(h_t)) u_{t+1}^{\pi}(h_t, d_t(h_t), j) \tag{2.4}$$

The optimal value function is simply the supremum of the value function over all possible policies:

$$u_t^*(h_t) = \sup_{\pi \in \prod^{HR}} u_t^{\pi}(h_t).$$

Using Lemma 2.1, Richard Bellman proposed a couple of equations whose solution would be the optimal value function.

$$\begin{aligned} u_t(h_t) &= \sup_{a \in A_{s_t}} \left\{ r_t(s_t, a) + \sum_{j \in S} p(j|s_t, a) u_{t+1}(h_t, a, j) \right\} \\ u_N(h_N) &= r_N(s_N). \end{aligned} \tag{2.5}$$

We shall now outline the proof for the existence of a deterministic optimal policy for any finite horizon MDP, and then show that one can find this optimal policy by simply solving the Bellman Optimality equations in (2.5).

Theorem 2.2. *Suppose u_t is a solution to the Optimality equations* (2.5)*, then,*

(a) $u_t(h_t) = u_t^*(h_t)$ for all $h_t \in H_t$, $t \in T$.

(b) $u_1(s_1) = v_N^*(s_1)$ for all $s_1 \in S$.

Proof. By backward induction, we show $u_n(h_n) \geq u_n^*(h_n)$ for all $n \in T$. For $n = N$, obviously $u_N(h_N) = r_N(h_N) = u_N^*(h_N)$.
Assume for all histories $h_t : u_t(h_t) \geq u_t^*(h_t)$ for $t = n+1, \cdots, N$. Let $\pi' = (d_1', \cdots, d_{N-1}')$ be an arbitrary policy in $\prod^{HR}$. For $t = n$, the optimality equation is :

$$
\begin{aligned}
u_n(h_n) &= \sup_{a \in A_{s_n}} \left\{ r_n(s_n, a) + \sum_{j \in S} p_n(j|s_n, a) u_{n+1}(h_n, a, j) \right\} \\
&\geq \sup_{a \in A_{s_n}} \left\{ r_n(s_n, a) + \sum_{j \in S} p_n(j|s_n, a) u_{n+1}^*(h_n, a, j) \right\}, \\
&\geq \sum_{a \in A_{s_n}} q_{d_n'(h_n)}(a) \left\{ r_n(s_n, a) + \sum_{j \in S} p_n(j|s_n, a) u_{n+1}^*(h_n, a, j) \right\} \\
&= u_n^{\pi'}(h_n).
\end{aligned}
$$

The other way inequality is shown by constructing a policy π' and showing that it has a value arbitrarily close to $u_n(h_n)$.
Since π' is an arbitrary policy and both way inequalities are true, we will have that $u_t(h_t) = u_t^*(h_t)$ for all $t \in T$.
(b) follows from the definition of $u_t(h_t)$ in Definition 2.1.5 for $t = 1$. □

Hence, we have shown that a solution to the optimality equations (2.5) would be the optimal value function.
The following theorems shows that we may search only among the deterministic policies to find an optimal policy. Theorem 2.3 states that if there exists an action

that attains the supremum in the right hand side of (2.5), then there is a history dependent deterministic policy that is optimal. The following Theorem 2.4 further proves that if an optimal deterministic history dependent policy exists, then there is also a Markovian deterministic policy that is optimal.

Theorem 2.3. *Let u_t^* be a solution of the Optimality equations* (2.5)*, and suppose that for each t and $s_t \in S$, there exists an $a'_{s_t} \in A_{s_t}$ for which*

$$\begin{aligned} & r_t(s_t, a'_{s_t}) + \sum_{j \in S} p_t(j|s_t, a'_{s_t}) u_{t+1}^*(h_t, a'_{s_t}, j) \\ &= \sup_{a \in A_{s_t}} \left\{ r_t(s_t, a) + \sum_{j \in S} p_t(j|s_t, a) u_{t+1}^*(h_t, a, j) \right\} \end{aligned} \tag{2.6}$$

for all $h_t \in H_t$. Then there exists a history dependent deterministic policy which is optimal.

Theorem 2.4. *Let u_t^* be a solution of the Optimality equations* (2.5)*, Then*

(a) For each $t \in T$, $u_t^*(h_t)$ depends on h_t only through s_t.

(b) If there exists an $a' \in A_{s_t}$ such that (2.6) *holds for each $s_t \in S$ and $t \in T$, then there exists an optimal policy which is deterministic and Markovian.*

Theorem 2.3 and Theorem 2.4 are proved in Section 3, Chapter 4 of Puterman [1]. Supremum over the action set A_s is attained when A_s is :

(a) Finite, and $r_t(s, a)$, $p_t(j|s, a)$ are continuous in a [The case for the Prefetching problem],

(b) Compact, and $r_t(s, a)$ is upper semi continuous in a for every s, t, there exists an $M < \infty$ such that $|r_t(s, a)| \leq M$, and $p_t(j|s, a)$ is lower semi continuous in a.

Thus, we only need to search among Markovian deterministic policies for an optimal policy and need not search among the larger set of randomized history dependent policies. Moreover, we can solve the Bellman Equations (2.5) to obtain such an optimal policy.

2.2 Infinite Horizon Average Reward Criterion

We move to the case where $N = \infty$. In this criterion, a simple summation of rewards would not yield any useful information about the policy, since it will be infinite most of the time. Hence we shall define the average reward of a policy as :

$$\phi_\pi(s) = \lim_{n\to\infty} \frac{E^\pi[\sum_{t=0}^{n} r_t(s_t, a_t)]}{n+1}. \tag{2.7}$$

If the limit does not exist, we shall define $\phi_\pi(s)$ as the $\liminf$. For all practical purposes, the limit exists.
π^* is the average reward optimal policy if :

$$\phi_{\pi^*}(s) = \max_\pi \phi_\pi(s). \tag{2.8}$$

Now we shall determine the conditions for such an optimal π^* to exist.

Theorem 2.5. *If there exists a bounded function $f(s)$ for every $s \in S$ and a constant g such that*

$$g + f(s) = \max_a \left[r(s,a) + \sum_{s'\in S} P(s,a,s') f(s') \right] \tag{2.9}$$

then there exists a π^* such that

$$g = \lim_\pi \phi_\pi(s) = \phi_{\pi^*}(s). \tag{2.10}$$

Proof. Let $h_t = (s_0, a_0, \cdots, s_t)$. Since $E[f(s_t)] = E[E[f(s_t)|h_{t-1}]]$, it follows that for any policy,

$$E^\pi \left[\sum_{t=1}^{n} [f(s_t) - E^\pi[f(s_t)|h_{t-1}]] \right] = 0 \tag{2.11}$$

We also have that :

$$\begin{aligned} E^\pi[f(s_t)|h_{t-1}] &= \sum_{s'} P(s_{t-1}, a_{t-1}, s') f(s') \\ &= r_{t-1}(s_{t-1}, a_{t-1}) + \sum_{s'} P(s_{t-1}, a_{t-1}, s') f(s') - r_{t-1}(s_{t-1}, a_{t-1}) \end{aligned}$$

$$\leq \max_a \left[r_{t-1}(s_{t-1}, a) + \sum_{s'} P(s_{t-1}, a, s') f(s') \right] - r_{t-1}(s_{t-1}, a_{t-1})$$

$$= g + f(s_{t-1}) - r_{t-1}(s_{t-1}, a_{t-1}) \tag{2.12}$$

There is equality for π^* since it is defined to take the maximizing action. Hence,

$$0 \geq E^{\pi} \left[\sum_{t=1}^{n} [f(s_t) - g - f(s_{t-1}) + r_{t-1}(s_{t-1}, a_{t-1})] \right] \tag{2.13}$$

which can be rearranged to obtain :

$$g \geq E^{\pi} \left[\frac{f(s_n)}{n} \right] - E^{\pi} \left[\frac{f(s_0)}{n} \right] + E^{\pi} \left[\frac{\sum_{t=1}^{n} r_{t-1}(s_{t-1}, a_{t-1})}{n} \right] \tag{2.14}$$

where equality is for π^*. If we allow $n \to \infty$ and use the fact that f is bounded, we would have that :

$$g \geq \phi_{\pi}(s_0) \tag{2.15}$$

with equality for π^* for all possible values of s_0. $\square$

Thus, if the conditions of Theorem 2.5 are satisfied, then a stationary optimal policy exists and may be found from the equation in (2.9).
Now we shall determine when the conditions in Theorem 2.5 hold true.
Denote by $v^{\alpha}(s)$ the optimal expected α discounted return function,

$$v^{\alpha}(s) = \max_a \left[r(s, a) + \alpha \sum_{s'} P(s, a, s') v^{\alpha}(s') \right]. \tag{2.16}$$

Merely taking the limit $\alpha \to 1$ and then maximizing the equation would not provide anything useful, as the limit is often infinite for all actions.
We shall assume that the state space is countable. Hence, we can order the sets with integer labels - $0, 1, \cdots$. Let us fix a state, call it s_f (fixed state). Now we define what we shall refer to as the α return of state s relative to s_f:

$$f^{\alpha}(s) = v^{\alpha}(s) - v^{\alpha}(s_f) \tag{2.17}$$

We shall use this in (2.16) to obtain :

$$(1 - \alpha) v^{\alpha}(s_f) + f^{\alpha}(s) = \max_a \left[r(s, a) + \alpha \sum_{s'} P(s, a, s') f^{\alpha}(s') \right]. \tag{2.18}$$

If we have some sequence $\alpha_n \to 1$, $f^{\alpha_n}(s') \to f(s')$, and $(1-\alpha_n)v^{\alpha_n}(s_f) \to g$, then we would obtain from (2.16),

$$g + f(s) = \max_a \left[r(s,a) + \alpha \sum_{s'} P(s,a,s') f^{\alpha}(s') \right]. \tag{2.19}$$

The summation and limit is interchangeable since f is bounded and the reward functions are also finite.
We analyse the above sequences in detail.

Theorem 2.6. *If there exists an $N < \infty$ such that $v^{\alpha}(s) - v^{\alpha}(s_f) < N$ for all α, s, then :*

- *There exists a bounded function $f(s)$ and a constant g satisfying (2.9);*
- *For some sequence $\alpha_n \to 1$, $f(s) = \lim_{n\to\infty}[v^{\alpha_n}(s) - v^{\alpha_n}(s_f)]$;*
- $\lim_{\alpha\to 1}(1-\alpha)v^{\alpha}(s_f) = g$.

Proof. Since the state space is countable, label the states as $0 = s_f, 1, 2, \cdots, n, \cdots$.
$f^{\alpha}(t)$ is uniformly bounded for every t, α by assumption of the theorem. Every bounded sequence contains a convergent subsequence.
Thus, $\alpha_{1,n} \to 1$ such that $\lim_{n\to\infty} f^{\alpha_{1,n}}(1) = f(1)$ exists.
Since $f^{\alpha_{1,n}}(2)$ is bounded, there is a subsequence $\{\alpha_{2,n}\}$ of $\{\alpha_{1,n}\}$ such that $\lim_{n\to\infty} f^{\alpha_{2,n}}(2) = f(2)$ exists.
In this manner $f^{\alpha_{n,n}}(i) \to f(i)$ for each i.
Since rewards are bounded, $(1-\alpha_n)v^{\alpha_n}(0)$ is bounded, and thus there is a subsequence $\{\alpha_{\bar{n}}\}$ of $\{\alpha_n\}$ for which $\lim_{n\to\infty}(1-\alpha_{\bar{n}})v^{\alpha_{\bar{n}}}(0) = g$ exists.
From (2.18), we would get that :

$$(1-\alpha_{\bar{n}})v^{\alpha}(0) + f^{\alpha_{\bar{n}}}(i) = \max_a \left[r(i,a) + \alpha_{\bar{n}} \sum_{j=0}^{\infty} P(i,a,j) f^{\alpha_{\bar{n}}}(j) \right]. \tag{2.20}$$

By taking $n \to \infty$, and the boundedness of $f^{\alpha_{\bar{n}}}(j)$, we can use the Lebesgue-bounded convergence theorem to get :

$$\sum_{j=0}^{\infty} P(i,a,j) f^{\alpha_{\bar{n}}(j)} \to \sum_{j=0}^{\infty} P(i,a,j) f(j) \text{ as } n \to \infty. \tag{2.21}$$

Since $(1-\alpha)v^\alpha(0)$ is bounded, for any sequence $\alpha_n \to 1$, there is a subsequence $\{\alpha_{\sigma(n)}, n \geq 1\}$ such that

$$\lim_{n\to\infty}(1-\alpha_{\sigma(n)})v^{\alpha_{\sigma(n)}}(0) \text{ exists}$$

and this limit must be g. Since the subsequence limit is g, it must follow that,

$$\lim_{n\to\infty}(1-\alpha_n)v^{\alpha_n}(0) = g. \tag{2.22}$$

$\square$

Finally, we shall look at a sufficient condition for $v^\alpha(i) - v^\alpha(0)$ to be uniformly bounded.

Theorem 2.7. *If for some state (call it 0), there is a constant $N < \infty$ such that $m_{i0}(\pi^\alpha) < N$ for all i, α, and the rewards are bounded by M, then $v^\alpha(i) - v^\alpha(0)$ is uniformly bounded, where $m_{i0}(\pi^\alpha)$ is the expected time to go from state i to state 0 when using the α discounted optimal policy π^α.*

Proof. Let $T = \min\{n : s_n = 0\}$. Then,

$$\begin{aligned}
v^\alpha(i) &= E_i^{\pi^\alpha}\left[\sum_{n=0}^{T-1}\alpha^n r(s_n, a_n)\right] + E_i^{\pi^\alpha}\left[\sum_{n=T}^{\infty}\alpha^n r(s_n, a_n)\right] \\
&\leq M E_i^{\pi^\alpha}(T) + v^\alpha(0)E_i^{\pi^\alpha}(\alpha^T) \\
&\leq MN + v^\alpha(0)
\end{aligned} \tag{2.23}$$

where M is the bound on rewards. We have used the strong Markov property with T being the stopping time to obtain the inequality.
Note that :

$$v^\alpha(i) \geq v^\alpha(0)E_i^{\pi^\alpha}(\alpha^T)$$

or,

$$v^\alpha(0) \leq v^\alpha(i) + [1 - E_i^{\pi^\alpha}(\alpha^T)]v^\alpha(0).$$

Thus $v^\alpha(0) \leq M/(1-\alpha)$, and $E(\alpha^T) \geq \alpha^{E(T)} \geq \alpha^N$ by Jensen's inequality.
Hence,

$$v^\alpha(0) \leq v^\alpha(i) + (1-\alpha^N)\cdot\frac{M}{1-\alpha} \leq v^\alpha(i) + MN \tag{2.24}$$

Thus we have the uniform boundedness property. □

Remark. If the state space is finite, and if every stationary policy gives rise to an irreducible Markov chain, then $v^\alpha(i) - v^\alpha(0)$ is uniformly bounded.

This concludes the outline for the existence of a stationary optimal policy for the infinite horizon average costs criterion.

Chapter 3

Prefetching MDP Specification

3.1 Design of the Prefetching MDP

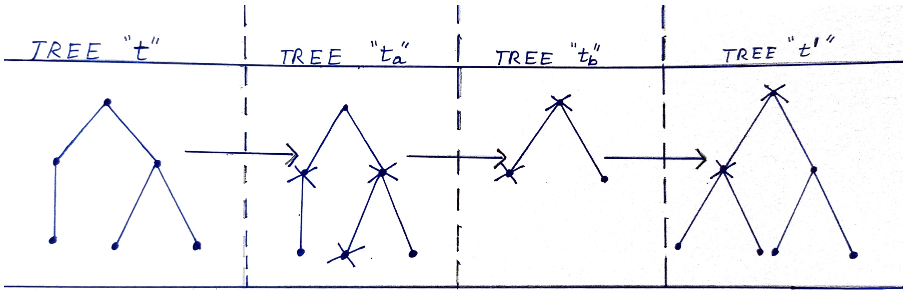

Figure 3.1: Flow of the Prefetching MDP

Figure 3.1 depicts the process of prefetching designed in the following manner :

- The surfer is on a rooted tree (t) of a specified depth parameter d.
- Each node of the tree, except the leaves, has between 1 and p sons. The parameter p is also a specified parameter of the model.
- Each node is either "marked" or "unmarked" to denote prefetched or non-prefetched information respectively. Marks are represented as crosses in Figure 3.1. The tree t_a is the tree after markings.

- The surfer moves from the root to one of the sons with uniform probability.
- The surfer is **not** allowed to move backwards, and is allowed to move to only one of the sons of the root.
- After the surfer moves to one of the sons, we focus only on the offspring of the particular node that the surfer is on.
- The surfer is now on the root of a depth $d-1$ tree t_b. We discover one depth further by attaching between 1 and p nodes to each leaf of the depth $d-1$ tree to obtain a depth d tree, t'.
- Assume that all possible combinations of forming a depth d tree from a given $d-1$ depth tree are equally likely (Uniform probability).

3.2 State Space

The state space S will be the set of all rooted and marked trees of depth d with nodes having sons between 1 and p. We use the notation $\mathcal{T}_{p,d}$ to denote that set.

$$S = \mathcal{T}_{p,d}.$$

For a finite p, d, S is a finite discrete set.
Representation of elements :
$t \in \mathcal{T}_{p,d}$ for $d > 0$ is represented as $t = (\mu, s)$ where $\mu \in \{0,1\}$ and $s \in \mathcal{L}_{d-1}$ where $\mathcal{L}_{d-1} = \mathcal{T}_{p,d-1} \cup (\mathcal{T}_{p,d-1} \times \mathcal{T}_{p,d-1}) \cup \ldots \cup (\mathcal{T}_{p,d-1})^p$ represents a list of subtrees.
For $d = 0$, we have $t = (\mu)$ where $\mu \in \{0,1\}$.
Let us denote the number of rooted trees of depth d with m nodes as $X_d(m)$.

Minimum number of nodes for a tree of depth d will be $d+1$.
Maximum number of nodes for a tree of depth d will be $(p^{d+1}-1)/(p-1)$.
Denote the bounds on number of nodes of depth d tree as minnode(d), and maxnode(d). Using this notation we will obtain,

$$|S| = \sum_{m=\text{minnode}(d)}^{\text{maxnode}(d)} 2^m \times X_d(m). \tag{3.1}$$

Since there are 2^m possible markings for a tree with m nodes.

Remark. The states in S which have any of the leaves marked never feature in our MDP. This is by definition of our process which involves discovery of unmarked leaves after the surfer moves.

Given a budget or a specific family of policies, there are states in S which will never feature in the MDP. Thus, we shall focus only on the "Usable States".

Definition 3.2.1 (Usable States U)**.** The states in S which are attained through transitions given a specific case of budget, or a family of policies are called usable states. Denote this set of states as U.

For example, budget dependent states for $k = 1, p = 2, d = 2$ will not include states where both nodes at depth 1 (if two exist) are marked.
We provide a recursion formula for $X_d(m)$.

$$\begin{aligned} X_d(m) =& X_{d-1}(m-1) + \sum_{\substack{m_1+m_2=m-1 \\ m_1,m_2\geq 1}} X_{d-1}(m_1)X_{d-1}(m_2) \\ &+ \sum_{\substack{m_1+m_2+m_3=m-1 \\ m_1,m_2,m_3\geq 1}} X_{d-1}(m_1)X_{d-1}(m_2)X_{d-1}(m_3) + \cdots \\ &+ \sum_{\substack{m_1+m_2+\cdots+m_{\min(p,m-1)}=m-1 \\ m_1,m_2,m_3,\cdots,m_{\min(p,m-1)}\geq 1}} X_{d-1}(m_1)X_{d-1}(m_2)\cdots X_{d-1}(m_{\min(p,m-1)}) \end{aligned} \tag{3.2}$$

To get an idea of the range of values of $X_d(m)$, we obtain an upper bound through the following computation.
Let $\gamma = \min(p-1, m-2)$, then the number of terms on the right hand side of 3.2 is bounded by :

$$\begin{aligned} \sum_{r=1}^{\min(p,m-1)} \binom{m-2}{r-1} &= \sum_{r=0}^{\min(p-1,m-2)} \binom{m-2}{r} \\ &\leq \sum_{r=0}^{\gamma} \frac{(m-2)^r}{r!} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{r=0}^{\gamma} \frac{\gamma^r * (\frac{m-2}{\gamma})^r}{r!} \\
&\le e^{\gamma} \left(\frac{m-2}{\gamma} \right)^{\gamma} \quad \forall p \ge 1.
\end{aligned} \tag{3.3}
$$

We found an expression for $X_2(m)$ as :

Proposition 3.1. *Assuming that* $\binom{a}{b} = 0$ *for all* $a \le 0$ *or* $a < b$ *:*

$$
X_2(m) = \sum_{r=1}^{\min(p,m-1)} \sum_{j=0}^{r} \binom{m-(r+2)-jp}{r-1} \times (-1)^j \times \binom{r}{j} \tag{3.4}
$$

We should note here that the size of S grows massively as we go from depth 2 to depth 3, or as we increase the parameter p for a particular depth.
To get an idea of the sizes of S for different model parameters, we calculate $|S|$ for a few cases in Table 3.1. Following this, in Table 3.2, for depth 2 trees we have explicitly mentioned the contribution from the terms $X_2(m)$ for possible values of m given $p = 1, 2, 3$.

Table 3.1: Size of State space for different d and p

$d \backslash p$	1	2	3	4
1	4	12	28	60
2	8	312	45,528	
3	16	197,360		
4	32			

Table 3.2: Size of State space for $d = 2$

p	$X_2(3)$	$X_2(4)$	$X_2(5)$	$X_2(6)$	$X_2(7)$	$X_2(8)$	$X_2(9)$	$X_2(10)$	$X_2(11)$	$X_2(12)$	$X_2(13)$	Number of Trees	$\lvert\mathcal{T}_{p,2}\rvert$
1	1											1	8
2	1	1	1	2	1							6	312
3	1	1	2	2	4	5	7	7	6	3	1	39	45,528

3.3 Action Space

- The action in the prefetching MDP is marking of nodes given a budget k.
- At every time step, a set of k vertices from the tree is chosen to be marked.
- The Action space A should contain k subsets of all possible trees.
- If k is greater than the number of vertices of the tree, then the action will be the whole tree itself.

$$A = \text{Set of all } k \text{ subsets of vertices of trees in } \mathcal{T}_{p,d}. \tag{3.5}$$

The size of the action space will be the sum of the number of ways to choose k vertices from the largest tree in $\mathcal{T}_{p,d}$ and the number of trees of depth d with number of vertices ranging from minnode(d) to k. Formally,

$$|A| = \binom{\frac{p^{d+1}-1}{p-1} - 1}{k} + \sum_{m=d+1}^{k} X_d(m). \tag{3.6}$$

The first term represents choosing k vertices from the largest tree in $\mathcal{T}_{p,d}$. The second term is to account for those trees which have atmost k vertices.

Remark. The Action space contains "wasteful" actions such as marking already marked nodes, and also actions which mark nodes that do not exist for particular trees.

3.4 Transition Probability Structure

To define the Transition probabilities, we introduce two set mappings.

Definition 3.4.1. $\mathcal{D} : \mathcal{T}_{p,d-1} \to \mathcal{P}(\mathcal{T}_{p,d})$
This can be understood as the "discovery" mapping of trees of depth $d-1$ to the power set of trees of depth d.
$\mathcal{D}(t)$ is the set of trees in $\mathcal{T}_{p,d}$ obtained from the tree t by updating leaves of t according to the below rule:
Let l be a leaf (i.e, depth 0 tree) in t. Update $l = (\mu)$ as

$$(\mu) \to (\mu, nl) \text{ where } nl \in \{0\} \cup \{0,0\} \cup ... \cup \{0\}^p. \tag{3.7}$$

Definition 3.4.2. $\mathcal{SD} : \mathcal{T}_{p,d} \to \mathcal{P}(\mathcal{T}_{p,d})$ defined by

$$\mathcal{SD}((\mu, s)) = \bigsqcup_{t_s \in s} \mathcal{D}(t_s). \tag{3.8}$$

where the symbol $\sqcup$ refers to the disjoint union of sets. This mapping can be understood as the "subtree discovery" mapping of depth d trees to the power set of depth d trees.

Let us recall the process flow :

- Begin at a state $t = (\mu, s)$
- Perform action a to move to $t_a = (\mu^*, s^*)$, i.e, $a((\mu, s)) = (a(\mu), a(s)) = (\mu^*, s^*)$.
- Surfer moves to one of the sub-trees $t_b \in s^*$
- Discover one depth further of t_b to obtain new state t'.

Hence, we will have :

$$P(t, a, t') = \begin{cases} \dfrac{1}{|s||\mathcal{D}(t_b)|} & \text{if } t' \in \mathcal{D}(t_b) \text{ where } t_b \in a(s) \\ \\ 0 & \text{if } t' \notin \mathcal{SD}(a(t)). \end{cases} \tag{3.9}$$

where $\mathcal{D}$ and $\mathcal{SD}$ are as defined in Definition 3.4.1 and Definition 3.4.2.

3.5 Cost Function

The immediate cost for moving to state t' by choosing action a while in state t is

$$c(t, a, t') = \begin{cases} 0 & \text{if } \mu' = 1 \\ 1 & \text{if } \mu' = 0 \end{cases} \tag{3.10}$$

$$c(t, a) = \sum_{t' \in \mathcal{SD}(t)} P(t, a, t') c(t, a, t'). \tag{3.11}$$

Since the immediate costs depend on the state that the process transitions to, we shall use the expected cost of choosing action a when in state t.

3.6 Bellman Optimality Equations

We model the Prefetching process as a Finite Horizon MDP. Hence, let us look at the Bellman Optimality Equation for the Prefetching MDP given a finite horizon N.

$$
\begin{aligned}
&V^*(n,t) = \min_{a \in A} \left[c(t,a) + \sum_{t' \in S} P(t,a,t') V^*(n+1,t') \right] \text{ for } 0 \leq n < N \\
&V^*(N,t) = 0 \text{ , and} \\
&\Gamma_n^*(t) \in \operatorname*{argmin}_{a \in A} V^*(n,t) \text{ for all } n < N.
\end{aligned}
\tag{3.12}
$$

The Value function under any policy $\Gamma = (\Gamma_0, \cdots, \Gamma_{N-1})$ is :

$$
\begin{aligned}
&V_\Gamma(n,t) = c(t,\Gamma_n(t)) + \sum_{t' \in S} P(t,\Gamma_n(t),t') V_\Gamma(n+1,t') \text{ for all } 0 \leq n \leq N-1, \\
&V_\Gamma(N,t) = 0 \text{ for all } t \in S.
\end{aligned}
$$

Let us make a convenient change of notation to work forwards, rather than backwards.

$$
\begin{aligned}
W^*(n,t) &= V^*(N-n,t) \text{ for all } 0 \leq n \leq N, \\
\pi_n &= \Gamma_{N-1-n} \text{ for all } 0 \leq n \leq N-1.
\end{aligned}
\tag{3.13}
$$

Hence solving the equations in (3.12) is equivalent to solving :

$$
\begin{aligned}
&W^*(n,t) = \min_{a \in A} \left[c(t,a) + \sum_{t' \in S} P(t,a,t') W^*(n-1,t') \right] \text{ for } 0 < n \leq N \\
&W^*(0,t) = 0 \text{ , and} \\
&\pi_n^*(t) \in \operatorname*{argmin}_{a \in A} W^*(n+1,t) \text{ for all } 0 \leq n \leq N-1.
\end{aligned}
\tag{3.14}
$$

Chapter 4

Finite Horizon Depth 1 Trees

In this chapter, we analyse Depth 1 trees and compute the optimal policy for arbitrary p, k. We shall first look at an example, and then proceed to finding the optimal policy for the general case. Consider the model parameters $p = 2, k = 1$ and $d = 1$.

- $S = \{(0, \{0\}), (1, \{0\}), (1, \{1\}),$
  $(0, \{(0, 0)\}),$
  $(1, \{(0, 0)\}), (0, \{(1, 0)\}), (0, \{(0, 1)\}),$
  $(1, \{(1, 0)\}), (0, \{(1, 1)\}), (1, \{(0, 1)\}),$
  $(1, \{(1, 1)\})\}$
- According to Definition 3.2.1, the Usable state space will be
  $U = \{(0, \{0\}), (1, \{0\}), (0, \{(0, 0)\}), (1, \{(0, 0)\})\}$.
- Let us denote the trees in U as t_1, t_2, t_3, t_4 corresponding to
  $\{(0, \{0\}), (1, \{0\}), (0, \{(0, 0)\}), (1, \{(0, 0)\})$.
- Label the nodes of the tree from the root node and proceed numbering the depth 1 nodes from the left. The root node is numbered 0.
- The action space for the usable state space will be $A = \{\{0\}, \{1\}, \{2\}\}$ where the actions are the node labels to be marked.

We now proceed to finding the value function of all the states in U using the Bellman equations (3.14). We shall start with $N = 1$ and then move to arbitrary N.

For $N = 1$ The process stops after one time step. Let us apply (3.14) with $N = 1$.

$$W^*(0,t_1) = 0, W^*(0,t_2) = 0, W^*(0,t_3) = 0, W^*(0,t_4) = 0. \tag{4.1}$$

$$\begin{aligned}
W^*(1,t_1) &= \min(1+\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 0\ ,\ 0+\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 0) = \min(1,0) = 0\\
W^*(1,t_2) &= \min(0+\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 0) = \min(0) = 0\\
W^*(1,t_3) &= \min(1+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0\ ,\ \frac{1}{2}+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0\\
&\quad ,\ \frac{1}{2}+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0) = \min(1,\frac{1}{2},\frac{1}{2}) = \frac{1}{2}\\
W^*(1,t_4) &= \min(\frac{1}{2}+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0\\
&\quad ,\ \frac{1}{2}+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0) = \min(\frac{1}{2},\frac{1}{2}) = \frac{1}{2}
\end{aligned} \tag{4.2}$$

Therefore, the corresponding optimal decision rules will be :

$$\pi_0^*(t_1) = \{1\}, \pi_0^*(t_2) = \{1\}, \pi_0^*(t_3) \in \{\{1\},\{2\}\}, \pi_0^*(t_4) \in \{\{1\},\{2\}\}. \tag{4.3}$$

For $N = 2$

$$\begin{aligned}
&W^*(0,t_1) = 0, W^*(0,t_2) = 0, W^*(0,t_3) = 0, W^*(0,t_4) = 0\\
&W^*(1,t_1) = 0, W^*(1,t_2) = 0, W^*(1,t_3) = 1/2, W^*(1,t_4) = 1/2\\
&W^*(2,t_1) = 1/4, W^*(2,t_2) = 1/4, W^*(2,t_3) = 3/4, W^*(2,t_4) = 3/4.
\end{aligned} \tag{4.4}$$

Therefore, corresponding optimal decision rules for $n = 1$ are:

$$\pi_1^*(t_1) = \{1\}, \pi_1^*(t_2) = \{1\}, \pi_1^*(t_3) \in \{\{1\},\{2\}\}, \pi_1^*(t_4) \in \{\{1\},\{2\}\}. \tag{4.5}$$

For $N = 3$

$$\begin{aligned}
&W^*(0,t_1) = 0, W^*(0,t_2) = 0, W^*(0,t_3) = 0, W^*(0,t_4) = 0\\
&W^*(1,t_1) = 0, W^*(1,t_2) = 0, W^*(1,t_3) = 1/2, W^*(1,t_4) = 1/2\\
&W^*(2,t_1) = 1/4, W^*(2,t_2) = 1/4, W^*(2,t_3) = 3/4, V^*(1,t_4) = 3/4\\
&W^*(3,t_1) = 1/2, W^*(3,t_2) = 1/2, W^*(3,t_3) = 1, W^*(3,t_4) = 1
\end{aligned} \tag{4.6}$$

The corresponding optimal decision rules are same as the $N = 2$ case for all values of $0 \leq n \leq 3$. This is a stationary policy.
We now seek to calculate the expected cost of the MDP. We start by calculating the stationary distribution of the Markov chain, and then the expected cost of transitions. Finally, we obtain the average cost of the process.

- Transition matrix over the four states if the optimal policy is followed:

$$P = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \tag{4.7}$$

- Stationary distribution :

$$\text{using } \rho = \rho \times P \text{ ,we get } \rho = \begin{pmatrix} 1/8 & 3/8 & 1/8 & 3/8 \end{pmatrix} \tag{4.8}$$

- Expected costs of transitions

$$E[c(t_1, a_1)] = 0, E[c(t_2, a_1)] = 0, E[c(t_3, a_1)] = 1/2, E[c(t_4, a_1)] = 1/2 \tag{4.9}$$

- Average cost :

$$1/8 \cdot 0 + 3/8 \cdot 0 + 1/8 \cdot 1/2 + 3/8 \cdot 1/2 = 1/4 \tag{4.10}$$

Remark. Similar calculations for the case $p = 3, d = 1, k = 2$ resulted in an average cost of 1/9.

4.1 Optimal Policy for the General Case

Let us consider the general case of depth 1 trees for arbitrary p and k.

- Label the states in U as follows:

$$\begin{aligned} &t_{0,1}, t_{0,2}, \cdots, t_{0,p}, \\ &t_{1,1}, t_{1,2}, \cdots, t_{1,p}. \end{aligned} \tag{4.11}$$

  where $t_{i,j}$ denotes the tree with j sons and i is the mark of root.

- Given an unmarked tree with j unmarked sons, where $1 \leq j \leq p$, the number of possible actions are $\binom{j}{k} + \binom{j}{k-1}$.
- Let us label the actions as below. The action space is union of the two sets:

$$\begin{aligned} A_1 &= \{a_1, a_2, \cdots, a_{\binom{j}{k}}\} \\ A_2 &= \{a_{\binom{j}{k}+1}, \cdots, a_{\binom{j}{k}+\binom{j}{k-1}}\}. \end{aligned} \tag{4.12}$$

  The Actions in A_1 represent every possible combination of marking sons. The Actions in A_2 represents every possible combination of marking the root and the sons.
- The cost function is :

$$c(t_{i,j}, \pi(t_{i,j})) = \frac{(j-k)^+}{j} \text{ for all } i \in \{0,1\}, 1 \leq j \leq p. \tag{4.13}$$

- The transition probability structure is :

$$P(t_{i,j}, \pi(t_{i,j}), t_{e,f}) = \begin{cases} \dfrac{(j-k)^+}{jp} & \text{if } e = 0, 1 \leq f \leq p \\ \dfrac{(j-(j-k)^+)}{jp} & \text{if } e = 1, 1 \leq f \leq p \end{cases} \tag{4.14}$$

Before getting to the formal proof of the optimal policy for depth one trees, let us prove a useful lemma.
We define the policy π to mark any k combination of sons at every decision epoch. Let $\tilde{\pi}$ be the policy that marks any k combination of sons upto $N-2$, and mark the root alongwith a $k-1$ combination of sons at $N-1$. Formally, using the notation in (4.12),

$$\begin{aligned} &\pi_n \in A_1 \text{ for all } 0 \leq n \leq N-1 \\ &\tilde{\pi}_n \in A_1 \text{ for all } 0 \leq n \leq N-2 \\ &\tilde{\pi}_{N-1} \in A_2. \end{aligned} \tag{4.15}$$

Lemma 4.1.

$$W_\pi(n, t_{0,j}) = W_\pi(n, t_{1,j}) \textit{ for all } 1 \leq j \leq p \text{ , } 0 \leq n \leq N.$$

Proof. We prove the lemma by recurrence. $W_\pi(0, t_{0,j}) = W_\pi(0, t_{1,j}) = 0$.
For $n = 1$:

$$W_\pi(1, t_{0,j}) = c(t_{0,j}, \pi_0(t_{0,j})) = \frac{(j-k)^+}{j}$$
$$W_\pi(1, t_{1,j}) = c(t_{1,j}, \pi_0(t_{1,j})) = \frac{(j-k)^+}{j}.$$

The statement holds true for $n = 1$. For any $2 \leq n \leq N$,

$$\begin{aligned}
W_\pi(n, t_{0,j}) &= c(t_{0,j}, \pi_{n-1}(t_{0,j})) + \sum_{t_{e,f} \in S} P(t_{0,j}, \pi_{n-1}(t_{0,j}), t_{e,f}) W_\pi(n-1, t_{e,f}) \\
&= c(t_{0,j}, \pi_{n-1}(t_{0,j})) + \sum_{t_{0,f} \in S} P(t_{0,j}, \pi_{n-1}(t_{0,j}), t_{0,f}) W_\pi(n-1, t_{0,f}) \\
&+ \sum_{t_{1,f} \in S} P(t_{0,j}, \pi_{n-1}(t_{0,j}), t_{1,f}) W_\pi(n-1, t_{1,f}) \qquad (4.16) \\
&= c(t_{0,j}, \pi_{n-1}(t_{0,j})) + \sum_{f=1}^{p} \left(\frac{(j-k)^+}{jp} W_\pi(n-1, t_{0,f}) \right. \\
&\left. + \frac{(j-(j-k)^+)}{jp} W_\pi(n-1, t_{1,f}) \right). \qquad (4.17)
\end{aligned}$$

In (4.16), we have split the summation over the trees that are marked and unmarked.
Since the transition probability of a tree $t_{i,j}$ does not depend on i, a similar calculation for $t_{1,j}$ results in:

$$\begin{aligned}
W_\pi(n, t_{1,j}) = c(t_{1,j}, \pi_{n-1}(t_{1,j})) + \sum_{f=1}^{p} &\left(\frac{(j-k)^+}{jp} W_\pi(n-1, t_{0,f}) \right. \\
&\left. + \frac{(j-(j-k)^+)}{jp} \; W_\pi(n-1, t_{1,f}) \right).
\end{aligned} \qquad (4.18)$$

From (4.13), the immediate costs are equal, and hence (4.17) and (4.18) are equal. □

Theorem 4.2. *Given $S = \mathcal{T}_{p,1}$, the optimal policy is marking any k - combination of sons for every $0 \leq n \leq N-1$.*

Proof. We use an Exchange Argument.

We shall look at the value of states at $n = N$ under the two policies $\pi, \tilde{\pi}$ as defined in (4.15), which will be:

$$W_\pi(N, t_{i,j}) = c(t_{i,j}, \pi_{N-1}(t_{i,j})) + \sum_{t_{e,f} \in S} P(t_{i,j}, \pi_{N-1}(t_{i,j}), t_{e,f}) W_\pi(N-1, t_{e,f})$$
$$\text{for all } 1 \le j, f \le p \text{ and } i, e \in \{0, 1\}, \tag{4.19}$$

$$W_{\tilde{\pi}}(N, t_{i,j}) = c(t_{i,j}, \tilde{\pi}_{N-1}(t_{i,j})) + \sum_{t_{e,f} \in S} P(t_{i,j}, \tilde{\pi}_{N-1}(t_{i,j}), t_{e,f}) W_{\tilde{\pi}}(N-1, t_{e,f})$$
$$\text{for all } 1 \le j, f \le p \text{ and } i, e \in \{0, 1\}. \tag{4.20}$$

The equations (4.19) and (4.20) will be the same for trees of kind $t_{1,j}$ since $\pi_{N-1}(t_{1,j}) = \tilde{\pi}_{N-1}(t_{1,j})$. Thus,

$$W_\pi(N, t_{1,j}) = W_{\tilde{\pi}}(N, t_{1,j}).$$

Hence, we focus only on the trees of kind $t_{0,j}$. The immediate costs under policy $\tilde{\pi}$ will be:

$$c(t_{0,j}, \tilde{\pi}_{N-1}(t_{0,j})) = \frac{(j-k+1)^+}{j}. \tag{4.21}$$

Clearly the cost of $t_{i,j}$ under Policy π is greater than the cost under the Policy $\tilde{\pi}$ for $j \ge k$, and equal for $j < k$.

The probability of transition to a tree that is unmarked is higher under the Policy $\tilde{\pi}$, specifically it is:

$$P(t_{0,j}, \tilde{\pi}_{N-1}(t_{0,j}), t_{e,f}) = \begin{cases} \dfrac{(j-k+1)^+}{jp} & \text{if } e = 0, 1 \le f \le p \\ \\ \dfrac{(j-(j-k+1)^+)}{jp} & \text{if } e = 1, 1 \le f \le p \end{cases} \tag{4.22}$$

Let us take difference between the second terms on RHS of (4.19) and (4.20).

$$\sum_{t_{e,f}\in S} P(t_{0,j},\tilde{\pi}_{N-1}(t_{0,j}),t_{e,f})W_{\tilde{\pi}}(N-1,t_{e,f}) - P(t_{0,j},\pi_{N-1}(t_{0,j}),t_{e,f})W_{\pi}(N-1,t_{e,f})$$
$$= \sum_{t_{e,f}\in S} \left(P(t_{0,j},\tilde{\pi}_{N-1}(t_{0,j}),t_{e,f}) - P(t_{0,j},\pi_{N-1}(t_{0,j}),t_{e,f})\right) W_{\pi}(N-1,t_{e,f}). \tag{4.23}$$

The above follows from the fact that until $N-2$, the policies have the same decision rules, hence the same values.

Let us split the sum and plug-in the value of transitions.

$$\sum_{t_{e,f}\in S} \left(P(t_{0,j},\tilde{\pi}_{N-1}(t_{0,j}),t_{e,f}) - P(t_{0,j},\pi_{N-1}(t_{0,j}),t_{e,f})\right) W_{\pi}(N-1,t_{e,f}) \tag{4.24}$$
$$= \sum_{f=1}^{p}\left(\left(\frac{(j-k+1)^+}{jp} - \frac{(j-k)^+}{jp}\right) W_{\pi}(N-1,t_{0,f})\right. \tag{4.25}$$
$$\left. + \left(\frac{j-(j-k+1)^+}{jp} - \frac{j-(j-k)^+}{jp}\right) W_{\pi}(N-1,t_{1,f})\right)$$
$$= \sum_{f=1}^{p}\left(\left(\frac{(j-k+1)^+ - (j-k)^+}{jp}\right) W_{\pi}(N-1,t_{0,f})\right.$$
$$\left. + \left(\frac{(j-k)^+ - (j-k+1)^+}{jp}\right) W_{\pi}(N-1,t_{1,f})\right) \tag{4.26}$$
$$= \sum_{f=1}^{p}\left(\frac{1}{jp}W_{\pi}(N-1,t_{0,f}) + \frac{-1}{jp}W_{\pi}(N-1,t_{1,f})\right) \text{ for } j \geq k,$$
$$= \sum_{f=1}^{p} W_{\pi}(N-1,t_{0,f})\left(\frac{1}{jp} - \frac{1}{jp}\right) \tag{4.27}$$
$$= 0.$$

Equation (4.26) is 0 for $j < k$. Equation (4.27) follows from Lemma 4.1.

The immediate costs under policy π is lesser than (or equal to) that under policy $\tilde{\pi}$ and the future costs are equal. Hence we have the following inequality:

$$W_{\pi}(N,t_{0,j}) \leq W_{\tilde{\pi}}(N,t_{0,j}). \tag{4.28}$$

where equality holds for the case $j < k$. Thus, the Proposed Policy π is better than Policy $\tilde{\pi}$. □

4.2 Proposed Policy Satisfies Optimality Equations

We seek to find the general expression for the value of a state under proposed policy π for arbitrary p, k, N. General recurrence relation for the value of a state under policy π for $0 < n \leq N$ is :

$$W_\pi(n, t_{i,j}) = \frac{(j-k)^+}{j} + \frac{1}{p}\sum_{f=1}^{p} W_\pi(n-1, t_{1,f}). \tag{4.29}$$

$$W_\pi(0, t_{i,j}) = 0 \text{ for all } 1 \leq j \leq p, i \in \{0,1\}. \tag{4.30}$$

For $n = 0, 1$, we simply have

$$W_\pi(1, t_{i,j}) = \frac{(j-k)^+}{j} \text{ for all } 1 \leq j \leq p, i \in \{0,1\}. \tag{4.31}$$

The following evaluations of $W_\pi(n, t_{i,j})$ for $2 \leq n \leq N$ hold for all trees $t_{i,j}$ with $1 \leq j \leq p, i \in \{0,1\}$.

$$\begin{aligned}
W_\pi(2, t_{i,j}) &= c(t_{i,j}, \pi(t_{i,j})) + \sum_{t_{e,f} \in S} P(t_{i,j}, \pi(t_{i,j}), t_{e,f}) W_\pi(1, t_{e,f}) \\
&= \frac{(j-k)^+}{j} + \sum_{t_{0,f} \in S} \frac{(j-k)^+}{jp} W_\pi(1, t_{0,f}) \\
&+ \sum_{t_{1,f} \in S} \frac{(j-(j-k)^+)}{jp} W_\pi(1, t_{1,f}) \\
&= \frac{(j-k)^+}{j} + \sum_{f=1}^{p} \frac{1}{p} W_\pi(1, t_{1,f}) \\
&= \frac{(j-k)^+}{j} + \frac{1}{p}\sum_{f=1}^{p} \frac{(f-k)^+}{f} \\
&= \frac{(j-k)^+}{j} + \frac{1}{p}\mathbb{H}_{pk}, \text{ where } \mathbb{H}_{pk} = \sum_{f=1}^{p} \frac{(f-k)^+}{f} \qquad (4.32)
\end{aligned}$$

We shall generously use the term $\mathbb{H}_{pk}$ as it simplifies calculations. The expansion of $\mathbb{H}_{pk}$ in terms of the harmonic series summation $\mathbb{H}_p$ for $k < p$ is :

$$
\begin{aligned}
\mathbb{H}_{pk} &= \frac{1}{p} \sum_{f=k+1}^{p} \frac{f-k}{f} \\
&= \frac{1}{p}(p-k-k \sum_{f=k+1}^{p} \frac{1}{f} \\
&= \frac{1}{p}\left(p-k-k(\mathbb{H}_p - \mathbb{H}_k)\right)
\end{aligned}
\tag{4.33}
$$

Equation (4.33) has the summation of harmonic series term $\mathbb{H}_p$ which does not have a closed form. From the Recurrence (4.29), let us calculate:

$$
\begin{aligned}
W_\pi(3, t_{i,j}) &= \frac{(j-k)^+}{j} + \frac{1}{p} \sum_{f=1}^{p} W_\pi(2, t_{1,f}) \\
&= \frac{(j-k)^+}{j} + \frac{1}{p} \sum_{f=1}^{p} \left(\frac{(f-k)^+}{f} + \frac{1}{p}\mathbb{H}_{pk} \right) \\
&= \frac{(j-k)^+}{j} + \frac{1}{p} \sum_{f=1}^{p} \frac{(f-k)^+}{f} + \frac{1}{p} \sum_{f=1}^{p} \frac{1}{p}\mathbb{H}_{pk} \\
&= \frac{(j-k)^+}{j} + \frac{1}{p}\mathbb{H}_{pk} + \frac{1}{p}\mathbb{H}_{pk} \\
&= \frac{(j-k)^+}{j} + \frac{2}{p}\mathbb{H}_{pk}.
\end{aligned}
\tag{4.34}
$$

Following the above pattern of calculations with a simple induction method and with the notation $\mathbb{H}_{pk}$ used in (4.32), we obtain the below Lemma for arbitrary n.

Lemma 4.3.

$$
W_\pi(n, t_{i,j}) = \frac{(j-k)^+}{j} + \frac{(n-1)}{p}\mathbb{H}_{pk} \textit{ for all } 1 \le n \le N. \tag{4.35}
$$

Similarly, for the policy $\tilde{\pi}$, we have the value equation for $n = N$:

$$
W_{\tilde{\pi}}(N, t_{i,j}) = \frac{(j-k+1)^+}{j} + \frac{(N-1)}{p}\mathbb{H}_{pk}. \tag{4.36}
$$

The action set is $A = \{a_1, \cdots, a_{\binom{j}{k}+\binom{j}{k-1}}\}$ which is the union of A_1, A_2 in (4.12). To calculate the minimum over A, we only have to check for two elements in A, one from each of the two sets $\{a_1, \cdots, a_{\binom{j}{k}}\}$ and $\{a_{\binom{j}{k}}, \cdots, a_{\binom{j}{k}+\binom{j}{k-1}}\}$.
This is because the actions in those two sets are just different combinations of marking and yield the same values.

For $1 \leq n \leq N$,

$$\begin{aligned}
W^*(n, t_{i,j}) &= \min_{a\in A}\left[c(t_{i,j}, a) + \sum_{t'\in S} P(t_{i,j}, a, t')W^*(n-1, t')\right] \\
&= \min\left\{\frac{(j-k)^+}{j} + \frac{(n-1)}{p}\mathbb{H}_{pk}, \frac{(j-k+1)^+}{j} + \frac{(n-1)}{p}\mathbb{H}_{pk}\right\} \\
&= \frac{(j-k)^+}{j} + \frac{(n-1)}{p}\mathbb{H}_{pk} \\
&= W_\pi(n, t_{i,j}), \text{ and} \\
\pi^*_{n-1}(t) &\in \{a_1, \cdots, a_{\binom{j}{k}}\} = \pi_{n-1}(t).
\end{aligned} \tag{4.37}$$

Hence, π is the optimal policy as the corresponding value function satisfies the Bellman Optimality equations (3.14).

Remark. The value of $W^*(n, t_{i,j})$ does not depend on the future t' that $t_{i,j}$ could transition to. Instead, it depends only on j and n.

Proposition 4.4. *The average cost of policy $\pi = \pi^*$ is $p^{-1}\mathbb{H}_{pk}$ for all $t_{i,j} \in S$.*

Proof.

$$\begin{aligned}
\lim_{N\to+\infty}\frac{W_\pi(N, t_{i,j})}{N} &= \lim_{N\to+\infty}\frac{1}{N}\left(\frac{(j-k)^+}{j} + \frac{(N-1)}{p}\mathbb{H}_{pk}\right) \\
&= \lim_{N\to+\infty}\left(\frac{(j-k)^+}{Nj} + \frac{(N-1)}{Np}\mathbb{H}_{pk}\right) \\
&= 0 + \lim_{N\to+\infty}\left(1 - \frac{1}{N}\right)\frac{1}{p}\mathbb{H}_{pk}, \\
&= \frac{1}{p}\mathbb{H}_{pk}.
\end{aligned}$$

□

Chapter 5

Finite Horizon Depth 2 Trees

In this chapter, we go one depth further to Depth two trees. We shall first study the case of budget one ($k = 1$) for an arbitrary p, and then move on to the more interesting case of budget two ($k = 2$).
We use the labels of trees of depth 1, $t_{i,j}$, to represent a depth 2 tree.
Hence, a tree of depth 2 can be represented as :

$$(\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\}), \tag{5.1}$$

where $1 \leq m \leq p, 1 \leq j_r \leq p, i_r \in \{0,1\}$ for all $1 \leq r \leq m$. This is in line with our representation of trees of depth d as a node attached to a list of depth $d-1$ trees.
We shall first look at a simple example. Consider the MDP parameters $p = 2, k = 1$ for $d = 2$. We shall first find the values of states in U under one particular policy. Let π be a policy that always marks a son of the root then the values $W_\pi(t)$ for $t \in U$ computed with equation (3.14) is in Table 5.1. We compare Table 5.1 with the value functions under the policy of marking a leaf, and infer that the policy of marking a son of the root is optimal.

Table 5.1: $W_\pi(n,t)$ for all trees in U for $0 \le n \le N = 4$

Tree	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$
$(\mu, \{t_{0,1}\})$	0	0	0	1/4	1/2
$(\mu, \{t_{0,2}\})$	0	0	1/2	3/4	1
$(\mu, \{t_{0,1}, t_{0,1}\})$	0	1/2	1/2	3/4	1
$(\mu, \{t_{0,2}, t_{0,1}\})$	0	1/2	3/4	1	5/4
$(\mu, \{t_{0,1}, t_{0,2}\})$	0	1/2	3/4	1	5/4
$(\mu, \{t_{0,2}, t_{0,2}\})$	0	1/2	1	5/4	3/2

Given $d = 2$ and an arbitrary p, we start the analysis of optimal policy for $k = 1$. Like in the depth 1 case, we shall use an exchange argument to prove that our proposed policy is optimal, and then show that our policy also satisfies the Bellman Optimality equations in (3.14), and hence is optimal.

5.1 Budget 1

Possible Policies $\bar{\pi}, \bar{\pi}, \tilde{\pi}$:

$$\begin{aligned} \pi &= (\pi_0, \pi_1, \cdots, \pi_{N-1}), \\ \bar{\pi} &= (\pi_0, \pi_1, \cdots, \bar{\pi}_{N-1}), \\ \tilde{\pi} &= (\pi_0, \pi_1, \cdots, \tilde{\pi}_{N-1}). \end{aligned} \tag{5.2}$$

where

$$\begin{aligned} \pi_n &= \text{Mark a son of the tree for all } 0 \le n \le N-1, \\ \bar{\pi}_{N-1} &= \text{ Mark a leaf of the tree,} \\ \tilde{\pi}_{N-1} &= \text{ Mark root of the tree.} \end{aligned} \tag{5.3}$$

We propose that the policy π is optimal. Let us compute the value function under the policy π.

Following π, the immediate cost for a tree t represented as $(\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\})$ where $1 \le m \le p$ will be

$$c(t, \pi(t)) = \frac{m-1}{m}. \tag{5.4}$$

The transition probability will be

$$P(t, \pi(t), t') = \frac{1}{m \cdot p^{j_r}} \text{ if } t' \in \mathcal{D}(t_{i_r, j_r}). \tag{5.5}$$

The value function for $n = 0, 1$ will trivially be :

$$\begin{aligned} &W_\pi(0, t) = 0, \\ &W_\pi(1, t) = c(t, \pi(t)) \text{ for all } 1 \leq m \leq p. \end{aligned}$$

For $n = 2$,

$$\begin{aligned} W_\pi(2, t) &= c(t, \pi(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t, \pi(t), t') \cdot W_\pi(1, t'), \\ &= 2 - \frac{1}{m} - \frac{1}{m} \sum_{r=1}^{m} \frac{1}{j_r}. \end{aligned} \tag{5.6}$$

The simplification is in Appendix A.1.
Moving one step further, let us compute the value function at $n = 3$.

$$\begin{aligned} W_\pi(3, t) &= c(t, \pi(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t, \pi(t), t') \cdot W_\pi(2, t') \\ &= 3 - \frac{1}{m} - \frac{1}{m} \cdot \sum_{r=1}^{m} \left(\frac{1}{j_r} + \mathbb{M}(1, j_r) \right) \end{aligned} \tag{5.7}$$

where

$$\mathbb{M}(1, j_r) = \frac{1}{p^j \cdot j} \sum_{1 \leq f_1, f_2, \cdots, f_j \leq p} \sum_{q=1}^{j} \frac{1}{f_q}. \tag{5.8}$$

Let us formally define the following recursion to aid our analysis. We shall solve the recursion later on.

$$\begin{aligned} \mathbb{M}(n, j) &= \frac{1}{p^j \cdot j} \sum_{1 \leq f_1, f_2, \cdots, f_j \leq p} \sum_{q=1}^{j} \left(\frac{1}{f_q} + \mathbb{M}(n-1, f_q) \right) \\ &\quad \text{for } 1 \leq n \leq N, \text{ and } 1 \leq j \leq p. \\ \mathbb{M}(0, j) &= 0 \text{ and } 1 \leq j \leq p. \end{aligned}$$

Following the computation for $W_\pi(3,t)$ we move from $n=3$ to $n=4$, to obtain,

$$\begin{aligned} W_\pi(4,t) &= 4 - \frac{1}{m} - \frac{1}{m}\sum_{r=1}^{m}\frac{1}{j_r} - \frac{1}{m}\cdot\sum_{r=1}^{m}\frac{1}{p^{j_r}j_r}\sum_{1\le f_1,f_2,\cdots,f_{j_r}\le p}\sum_{q=1}^{j_r}\left(\frac{1}{f_q} + \mathbb{M}(1,f_q)\right) \\ &= 4 - \frac{1}{m} - \frac{1}{m}\cdot\sum_{r=1}^{m}\left(\frac{1}{j_r} + \mathbb{M}(2,j_r)\right). \end{aligned} \tag{5.9}$$

Hence, by recurrence, the general equation for $W_\pi(n,t)$, for all $2 \le n \le N$ is:

$$W_\pi(n,t) = n - \frac{1}{m} - \frac{1}{m}\sum_{r=1}^{m}\left(\frac{1}{j_r} + \mathbb{M}(n-2,j_r)\right). \tag{5.10}$$

In order to simplify the value function and remove dependency on the term $\mathbb{M}(n,j)$, we solve the recurrence for $\mathbb{M}(n,j)$ (proof in Appendix A.2) to obtain :

$$\mathbb{M}(n,j) = \frac{n}{p}\cdot\mathbb{H}_p \tag{5.11}$$

We can use the simplified expression for $\mathbb{M}(n,j)$ in (5.10) which leads to the following Lemma 5.1 for $W_\pi(n,t)$.

Lemma 5.1.

$$W_\pi(n,t) = n - \frac{1}{m} - \frac{n-2}{p}\cdot\mathbb{H}_p - \frac{1}{m}\sum_{r=1}^{m}\frac{1}{j_r} \quad \textit{for } 2 \le n \le N. \tag{5.12}$$

We can cross-verify (5.12) with the values in Table 5.1. Consider the tree $t = (\mu, \{t_{0,1}, t_{0,1}\})$ and $n=3$.
According to (5.12),

$$W_\pi(3,t) = 3 - \frac{1}{2} - \frac{3-2}{2}\mathbb{H}_2 - \frac{1}{2}\{1+1\} = 2 - \frac{1}{2} - \frac{3}{4} = \frac{3}{4}, \tag{5.13}$$

which is the corresponding value in Table 5.1. All values computed using (5.12) match with Table 5.1.

5.1.1 Exchange Argument Approach

Similar to the depth one case, let us prove a couple of useful lemmas before computing the optimal policy.

Lemma 5.2. *For a tree* $(\mu, \{t_{1,f_1}, t_{0,f_2}, \cdots, t_{0,f_j}\})$*, with* $j > 1$ *and given policies* π *and* $\bar{\pi}$*,*

$$\begin{aligned} &W_{\bar{\pi}}(n, (\mu, \{t_{1,f_1}, t_{0,f_2}, \cdots, t_{0,f_j}\})) - W_{\pi}(n, (\mu, \{t_{0,f_1}, t_{0,f_2}, \cdots, t_{0,f_j}\})) \\ &= -\frac{1}{j} \text{ for all } 1 \leq n \leq N-1. \end{aligned} \tag{5.14}$$

Proof. Let $t_m = (\mu, \{t_{1,f_1}, t_{0,f_2}, \cdots, t_{0,f_j}\})$ and $t_{um} = (\mu, \{t_{0,f_1}, t_{0,f_2}, \cdots, t_{0,f_j}\})$

$$\begin{aligned} W_{\bar{\pi}}(n, t_m) &= c(t, \bar{\pi}_{n-1}(t_m)) + \sum_{t' \in \mathcal{SD}(t_m)} P(t, \bar{\pi}_{n-1}(t_m), t') \cdot W_{\bar{\pi}}(n-1, t') \\ &= 1 - \frac{2}{j} + \sum_{t' \in \mathcal{SD}(t_m)} P(t_m, \bar{\pi}_{n-1}(t_m), t') \cdot W_{\bar{\pi}}(n-1, t') \end{aligned} \tag{5.15}$$

$$\begin{aligned} W_{\pi}(n, t_{um}) &= c(t, \pi_{n-1}(t_{um})) + \sum_{t' \in \mathcal{SD}(t_{um})} P(t_{um}, \pi_{n-1}(t), t') \cdot W_{\pi}(n-1, t') \\ &= 1 - \frac{1}{j} + \sum_{t' \in \mathcal{SD}(t_{um})} P(t_{um}, \pi_{n-1}(t_{um}), t') \cdot W_{\pi}(n-1, t'). \end{aligned} \tag{5.16}$$

Taking the difference of (5.15) and (5.16) we get:

$$\begin{aligned} &W_{\bar{\pi}}(n, t_m) - W_{\pi}(n, (\mu, t_{um}) \\ &= \frac{-1}{j}. \end{aligned}$$

This is because $\pi_i = \bar{\pi}_i$ for all $0 \leq i \leq N-2$, and the values of discoveries of marked and unmarked tree of same shape are equal.
Precisely,

$$t' \in \mathcal{D}(t_{1,f_1}) \implies t' = (1, \{t_{0,h_1}, \cdots, t_{0,h_{f_1}}\}) \text{ where } 1 \leq h_1, \cdots, h_{f_1} \leq p. \tag{5.17}$$

For such a t', there is a corresponding unmarked tree of same shape :

$$t'' \in \mathcal{D}(t_{0,f_1}) \implies t'' = (0, \{t_{0,h_1}, \cdots, t_{0,h_{f_1}}\}) \text{ where } 1 \leq h_1, \cdots, h_{f_1} \leq p. \tag{5.18}$$

Since, policy π does not mark the root, the immediate costs and future trees are same for t', t'', i.e.,

$$W_{\bar{\pi}}(n, t') = W_{\pi}(n, t'') \text{ for all } 1 \leq n \leq N-1. \tag{5.19}$$

Furthermore, the transition probabilities are same as well since the shapes of the future trees are identical :

$$P(t, \tilde{\pi}_{n-1}(t), t') = P(t, \pi_{n-1}(t), t'') \text{ for all } 1 \leq n \leq N-1. \tag{5.20}$$

Hence, the summations in (5.15) and (5.16) are equal. $\square$

Lemma 5.3. *Let, for all $n \geq 0$, all $\mu \in \{0, 1\}$, and all pairs of trees $t_{1,1}$, $t_{0,1} \in \mathcal{T}_{p,1}$,*

$$A_n = W_{\tilde{\pi}}(n, (\mu, \{t_{1,1}\})) - W_{\pi}(n, (\mu, \{t_{0,1}\})).$$

Then,

$$A_n = \begin{cases} -\sum_{f=2}^{p} \left\{\frac{1}{f}\right\} \cdot \left(\frac{1}{p} + \cdots + \frac{1}{p^{n-2}}\right) & \text{if } 3 \leq n \leq N-1 \\ 0 & \text{if } n < 3. \end{cases}$$

Furthermore,

$$1 + A_n > 0 \text{ for all } 0 \leq n \leq N-1.$$

Proof. The case $n = 0$ is trivial.
For $n = 1$, we just have the immediate costs which is 0 for both policies.

$$W_{\tilde{\pi}}(1, (\mu, \{t_{1,1}\})) = W_{\pi}(1, (\mu, \{t_{0,1}\})) = 0. \tag{5.21}$$

For $n = 2$, we have the immediate zero costs and future costs of discoveries of a subtree with one son, which is also zero.

$$W_{\tilde{\pi}}(2, (\mu, \{t_{1,1}\})) = W_{\pi}(2, (\mu, \{t_{0,1}\})) = 0. \tag{5.22}$$

Before moving on to $n = 3$, let us look at for some $2 \leq f \leq p$,

$$\begin{aligned} & W_{\tilde{\pi}}(2, (\mu, \{t_{1,f}\})) - W_{\pi}(2, (\mu, \{t_{0,f}\})) \\ &= \frac{1}{p^f} \sum_{t' \in \mathcal{SD}((\mu, \{t_{1,f}\}))} W_{\tilde{\pi}}(1, t') - \frac{1}{p^f} \sum_{t' \in \mathcal{SD}((\mu, \{t_{0,f}\}))} W_{\pi}(1, t') \\ &= \frac{1}{p^f} \sum_{1 \leq q_1, \cdots, q_f \leq p} W_{\tilde{\pi}}(1, (\mu, \{t_{1,q_1}, t_{0,q_2}, \cdots, t_{0,q_f}\})) \end{aligned}$$

$$- \frac{1}{p^f} \sum_{1 \leq q_1, \cdots, q_f \leq p} W_\pi(1, (\mu, \{t_{0,q_1}, t_{0,q_2}, \cdots, t_{0,q_f}\}))$$
$$= \frac{-p^f}{p^f f} = \frac{-1}{f}. \tag{5.23}$$

In (5.23), we have used Lemma 5.14. Similarly, we obtain by recurrence that, for any $2 \leq n \leq N-1$ and $2 \leq f \leq p$:

$$\begin{aligned} & W_{\bar{\pi}}(n, (\mu, \{t_{1,f}\})) - W_\pi(n, (\mu, \{t_{0,f}\})) \\ & = \frac{1}{p^f} \sum_{t' \in \mathcal{SD}((\mu, \{t_{1,f}\}))} W_{\bar{\pi}}(n-1, t') - \frac{1}{p^f} \sum_{t' \in \mathcal{SD}((\mu, \{t_{0,f}\}))} W_\pi(n-1, t') \\ & = \frac{1}{p^f} \sum_{1 \leq q_1, \cdots, q_f \leq p} W_{\bar{\pi}}(n-1, (1, \{t_{1,q_1}, t_{0,q_2}, \cdots, t_{0,q_f}\})) \\ & - \frac{1}{p^f} \sum_{1 \leq q_1, \cdots, q_f \leq p} W_\pi(n-1, (1, \{t_{0,q_1}, t_{0,q_2}, \cdots, t_{0,q_f}\})) \\ & = \frac{-1}{f}. \end{aligned} \tag{5.24}$$

Now we shall look at $n = 3$:

$$\begin{aligned} A_3 &= W_{\bar{\pi}}(3, (\mu, \{t_{1,1}\})) - W_\pi(3, (\mu, \{t_{0,1}\})) \\ &= \frac{1}{p} \cdot \sum_{f=1}^{p} (W_{\bar{\pi}}(2, (\mu, \{t_{1,f}\})) - W_\pi(2, (\mu, \{t_{0,f}\}))) \\ &= \frac{1}{p} \cdot (W_{\bar{\pi}}(2, (\mu, \{t_{1,1}\})) - W_\pi(2, (\mu, \{t_{0,1}\}))) \\ &+ \frac{1}{p} \cdot \sum_{f=2}^{p} (W_{\bar{\pi}}(2, (\mu, \{t_{1,f}\})) - W_\pi(2, (\mu, \{t_{0,f}\}))) \\ &= \frac{1}{p} \cdot A_2 + \frac{1}{p} \cdot \sum_{f=2}^{p} -\frac{1}{f} \qquad (5.25) \\ &= \frac{1}{p} \cdot \sum_{f=2}^{p} -\frac{1}{f} = \frac{1}{p} \cdot (1 - \mathbb{H}_p). \qquad (5.26) \end{aligned}$$

In (5.25), we have used the simplified value from (5.24).

The difference in value of trees $(\mu, \{t_{1,1}\}), (\mu, \{t_{0,1}\})$ for $3 \leq n \leq N-1$ is equivalent to a recurrence where $A_n = W_{\bar{\pi}}(n, (\mu, \{t_{1,1}\})) - W_{\pi}(n, (\mu, \{t_{0,1}\}))$:

$$\begin{aligned}
A_n &= \frac{1}{p} \cdot \sum_{f=1}^{p} \left(W_{\bar{\pi}}(n-1, (\mu, \{t_{1,f}\})) - W_{\pi}(n-1, (\mu, \{t_{0,f}\}))\right) \\
&= \frac{1}{p} \cdot \left(W_{\bar{\pi}}(n-1, (\mu, \{t_{1,1}\})) - W_{\pi}(n-1, (\mu, \{t_{0,1}\}))\right) \\
&+ \frac{1}{p} \cdot \sum_{f=2}^{p} \left(W_{\bar{\pi}}(n-1, (\mu, \{t_{1,f}\})) - W_{\pi}(n-1, (\mu, \{t_{0,f}\}))\right) \\
&= \frac{1}{p} \cdot A_{n-1} + \frac{1}{p} \cdot \sum_{f=2}^{p} -\frac{1}{f} \\
&= \frac{1}{p} \cdot A_{n-1} + \frac{1}{p} \cdot (1 - \mathbb{H}_p).
\end{aligned}$$

Hence,

$$\begin{aligned}
&A_n = \frac{1}{p} \cdot A_{n-1} + \frac{1}{p} \cdot (1 - \mathbb{H}_p) \text{ with } 3 \le n \le N-1 \text{ given} \\
&A_2 = 0.
\end{aligned} \tag{5.27}$$

Solving the above recurrence, we get,

$$\begin{aligned}
A_n &= (1 - \mathbb{H}_p) \cdot \left(\frac{1}{p} + \frac{1}{p^2} + \cdots + \frac{1}{p^{n-2}}\right) \\
&= -\sum_{f=2}^{p} \frac{1}{f} \cdot \left(\frac{1}{p} + \frac{1}{p^2} + \cdots + \frac{1}{p^{N-3}}\right) \text{ for all } 3 \le n \le N-1.
\end{aligned} \tag{5.28}$$

Clearly A_n is negative since $\mathbb{H}_p \ge 1$. Let us bound the positive term $-A_n$ for any $p > 1$:

$$\begin{aligned}
-A_n &= (\mathbb{H}_p - 1) \cdot \left(\frac{1}{p} + \frac{1}{p^2} + \cdots + \frac{1}{p^{n-2}}\right) \\
&< (\mathbb{H}_p - 1) \cdot \left(\sum_{k=1}^{\infty} \frac{1}{p^k}\right) \\
&= (\mathbb{H}_p - 1) \cdot \frac{1}{p-1} \\
&\le \frac{\ln p}{p-1} \le 1
\end{aligned} \tag{5.29}$$

We have used the well known logarithmic bound for the harmonic series term in (5.29). Precisely, we have used that $\mathbb{H}_p \le \ln p + 1$ for any $p \ge 1$. We then use the bound $\ln p \leq p-1$ for any $p \geq 1$.

Therefore, $A_n > -1 \implies 1+A_n > 0$ which completes the proof of the Lemma. $\square$

Theorem 5.4. *Given the state space $S = \mathcal{T}_{p,2}$, budget $k=1$, and finite horizon N, the optimal policy is marking any depth 1 node.*

Proof. Using the depth two tree representation as in (5.1), the cost of tree t under the three distinct policies $\pi, \bar{\pi}, \tilde{\pi}$ as defined in (5.2) are :

$$\begin{aligned} c(t, \pi(t)) &= 1 - \frac{1}{m} \\ c(t, \bar{\pi}_{N-1}(t)) &= 1 \\ c(t, \tilde{\pi}_{N-1}(t)) &= 1. \end{aligned} \tag{5.30}$$

Clearly,

$$W_{\pi}(n,t) = W_{\bar{\pi}}(n,t) = W_{\tilde{\pi}}(n,t) \text{ for all } 0 \leq n \leq N-1. \tag{5.31}$$

Let us compare the policies $\pi, \bar{\pi}$.

After applying decision rule $\bar{\pi}_{N-1}$, there would be future states which have already marked node at depth 1. Hence, we have to account for the value of these states. Without loss of generality, we assume that the first son of tree t_{i_c,j_c} i.e., the leaf, is marked under $\bar{\pi}_{N-1}$ where $c \in \{1, \cdots, m\}$.

$$\begin{aligned} &W_{\bar{\pi}}(N,t) \\ &= 1 + \sum_{t' \in \mathcal{SD}(t)} P(t, \bar{\pi}_{N-1}(t), t') \cdot W_{\bar{\pi}}(N-1, t') \\ &= 1 + \sum_{r=1}^{m} \left(\frac{1}{m \cdot p^{j_r}} \sum_{t' \in \mathcal{D}(t_{i_r,j_r})} W_{\bar{\pi}}(N-1, t') \right) \\ &= 1 + \sum_{r \in \{1,\cdots,m\}/c} \left(\frac{1}{m \cdot p^{j_r}} \sum_{1 \leq f_1, f_2, \cdots, f_{j_r} \leq p} W_{\bar{\pi}}(N-1, (i_r, \{t_{0,f_1}, t_{0,f_2}, \cdots, t_{0,f_{j_r}}\})) \right) \end{aligned}$$

$$+\left(\frac{1}{m\cdot p^{j_c}}\sum_{1\le f_1,\cdots,f_{j_c}\le p} W_{\bar{\pi}}(N-1,(i_c,\{t_{1,f_1},t_{0,f_2},\cdots,t_{0,f_{j_c}}\}))\right). \tag{5.32}$$

And under the policy π, we would have :

$$\begin{aligned}
&W_\pi(N,t)\\
&=1-\frac{1}{m}+\sum_{t'\in\mathcal{SD}(t)} P(t,\pi_{N-1}(t),t')\cdot W_\pi(N-1,t')\\
&=1-\frac{1}{m}\\
&+\sum_{r\in\{1,\cdots,m\}/c}\left(\frac{1}{m\cdot p^{j_r}}\sum_{1\le f_1,f_2,\cdots,f_{j_r}\le p} W_\pi(N-1,(i_r,\{t_{0,f_1},t_{0,f_2},\cdots,t_{0,f_{j_r}}\}))\right)\\
&+\left(\frac{1}{m\cdot p^{j_c}}\sum_{1\le f_1,f_2,\cdots,f_{j_c}\le p} W_\pi(N-1,(i_c,\{t_{0,f_1},t_{0,f_2},\cdots,t_{0,f_{j_c}}\}))\right).
\end{aligned} \tag{5.33}$$

Now, let us take difference between the Value Equations (5.32), and (5.33). The summation of terms over $r\in\{1,\cdots,m\}/c$ in (5.33) and (5.32) cancel out when we take their difference. For the remaining term, two cases must be considered. Case 1: $j_c=1$, using Lemma 5.3 for the value of discoveries of the subtree t_{i_c,j_c},

$$\begin{aligned}
&W_{\bar{\pi}}(N,t)-W_\pi(N,t)\\
&=\frac{1}{m}+\frac{1}{m\cdot p^{j_c}}\cdot\sum_{1\le f_1,f_2,\cdots,f_{j_c}\le p}\left(-\sum_{f=2}^{p}\left\{\frac{1}{f}\right\}\cdot\left(\frac{1}{p}+\frac{1}{p^2}+\cdots+\frac{1}{p^{N-3}}\right)\right)\\
&=\frac{1}{m}-\frac{1}{m\cdot p^{j_c}}\cdot p^{j_c}\cdot\sum_{f=2}^{p}\left\{\frac{1}{f}\right\}\cdot\left(\frac{1}{p}+\frac{1}{p^2}+\cdots+\frac{1}{p^{N-3}}\right)\\
&=\frac{1}{m}-\frac{1}{m}\cdot\sum_{f=2}^{p}\left\{\frac{1}{f}\right\}\cdot\left(\frac{1}{p}+\frac{1}{p^2}+\cdots+\frac{1}{p^{N-3}}\right)\\
&=\frac{1}{m}\cdot(1+A_{N-1})>0.
\end{aligned} \tag{5.34}$$

Case 2: $j_c \geq 2$, using Lemma 5.14,

$$\begin{aligned}
&W_{\bar{\pi}}(N,t) - W_{\pi}(N,t) \\
&\quad = \frac{1}{m} + \frac{1}{mp^{j_c}} \cdot \sum_{1 \leq f_1, f_2, \cdots, f_{j_c} \leq p} \frac{-1}{j_c} \\
&\quad = \frac{1}{m} + \frac{1}{mp^{j_c}} \cdot \frac{-1}{j_c} \cdot p^{j_c} \\
&\quad = \frac{1}{m} - \frac{1}{m \cdot j_c} \\
&\quad = \frac{1}{m} \cdot (1 - \frac{1}{j_c}) \\
&\quad > 0 \text{ since } j_c \geq 2.
\end{aligned} \tag{5.35}$$

Hence in all cases,

$$W_{\pi}(N,t) < W_{\bar{\pi}}(N,t). \tag{5.36}$$

Under policy $\tilde{\pi}$, there would be a higher immediate cost. The future costs are equal to the future costs under policy π since we follow the same decision rule and have the same future trees. The value for unmarked trees under this policy is:

$$W_{\tilde{\pi}}(N,t) = 1 + \sum_{t' \in \mathcal{SD}(t)} P(t, \tilde{\pi}_{N-1}(t), t') \cdot W_{\tilde{\pi}}(N-1, t'). \tag{5.37}$$

Taking difference of (5.37) and (5.33), we get :

$$W_{\tilde{\pi}}(N,t) - W_{\pi}(N,t) = \frac{1}{m} > 0. \tag{5.38}$$

Hence,

$$W_{\pi}(N,t) < W_{\tilde{\pi}}(N,t) \text{ for } t = (0, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\}). \tag{5.39}$$

When the decision rule of the policies $\bar{\pi}, \tilde{\pi}$ at $N-1$ are exchanged with π_{N-1}, we have a better of policy.

Hence, by the principle of exchange arguments, π (marking a depth one node) is the optimal policy. □

5.1.2 Proposed Policy Satisfies Optimality Equations

Let us prove that the Value Equation (5.12) satisfies the Optimality equations (3.14). We note the following points :

- The action set is the union of set of all singleton depth one nodes and the set of all singleton leaves.
- Among the actions that mark a leaf, we consider two types of actions that result in different values. One is marking a leaf with no siblings, and another is marking a leaf with at-least one sibling.
- The three actions we compare below in order are - First Action : Marking a depth 1 node, Second Action : Marking a leaf with at-least one sibling, Third Action : Marking a leaf with no sibling. i.e, $A = \{$ depth one node, leaf with atleast one sibling, leaf with no siblings $\}$.

Let $t = (\mu, \{t_{i_1,j_1}, t_{i_2,j_2}, \cdots, t_{i_m,j_m}\})$ where $1 \leq m \leq p$. Then we have,
For $n = 0$:

$$W^*(0,t) = W_\pi(0,t). \tag{5.40}$$

For $n = 1$:

$$\begin{aligned} W^*(1,t) &= \min_{a \in A} \left\{ c(t,a) + \sum_{t' \in S} P(t,a,t') \times W^*(0,t') \right\} \\ &= \min \left\{ 1 - \frac{1}{m}, 1, 1 \right\} \\ &= 1 - \frac{1}{m} \\ &= W_\pi(1,t), \text{ and} \\ \pi_0^*(t) &\in d_1(t) = \pi_0(t). \end{aligned} \tag{5.41}$$

The calculations for the other two actions have not been shown since they are rather trivial computations. Given that the leaf of subtree t_{i_c,j_c} (where $j_c \geq 2$ in the second action, and $j_c = 1$ in the third action) is marked in the other two actions, the immediate costs are 1 of course. Furthermore, the future values of discoveries of subtrees t_{i_r,j_r} for $r \neq c$ are equal to the future values under the action of marking a son (the first action). It is only the future values of the discoveries of the subtree t_{i_c,j_c} that is different. In comparision to the first action, when $j_c \geq 2$, there is an extra $1/j_c$ factor multiplied with the $-1/m$ term. For $j_c = 1$, the $1/j_c$ term actually does not feature at all! This shows that it is wasteful to mark the the leaf of a subtree with no siblings, since in the next step, given one budget this leaf (which would become a son if the surfer moves to this particular subtree) can be easily marked.

For $n = 2$:

$$
\begin{aligned}
W^*(2,t) &= \min_{a\in A}\left\{c(t,a) + \sum_{t'\in S} P(t,a,t') \times W^*(1,t')\right\} \\
&= \min\left\{2 - \frac{1}{m} - \frac{1}{m}\sum_{r=1}^{m}\frac{1}{j_r},\right. \\
&\qquad 2 - \frac{1}{m}\left(\frac{2}{j_c}\right) - \frac{1}{m}\sum_{r\in\{1,\cdots,m\}/c}\frac{1}{j_r}, \\
&\qquad \left. 2 - \frac{1}{m} - \frac{1}{m}\sum_{r\in\{1,\cdots,m\}/c}\frac{1}{j_r}\right\} \\
&\qquad = 2 - \frac{1}{m} - \frac{1}{m}\sum_{r=1}^{m}\frac{1}{j_r} = W_\pi(2,t), \text{ and} \\
&\qquad \pi_1^*(t) \in d_1(t) = \pi_1(t).
\end{aligned}
\tag{5.42}
$$

In (5.42), the minimum is the first action since the difference between the first and second action is a factor of $2/j_c$ multiplied to $-1/m$. Since $j_c \geq 2$, the factor is less than 1, and hence the negative terms in the second action are lesser in absolute value than the negative terms in the first action. The third action has an entire $-1/j_c = -1$ term missing and hence is not the minimum. The same argument applies in the following case for $2 < n \leq N$ in (5.43).

For $2 < n \leq N$

$$\begin{aligned}
W^*(n,t) = \min_{a\in A}&\left\{c(t,a) + \sum_{t'\in S} P(t,a,t') \times W^*(n-1,t')\right\}\\
= \min&\left\{n - \frac{1}{m} - \frac{1}{m}\cdot\sum_{r=1}^{m}\left(\frac{1}{j_r} + \mathbb{M}(n-2,j_r)\right),\right.\\
&n - \frac{1}{m}\cdot\left(\left(\frac{2}{j_c}\right) + \mathbb{M}(n-2,j_c)\right)\\
&-\frac{1}{m}\cdot\sum_{r\in\{1,\cdots,m\}/c}\left(\frac{1}{j_r} + \mathbb{M}(n-2,j_r)\right),\\
&\left.n - \frac{1}{m}\cdot\mathbb{M}(n-2,1) - \frac{1}{m}\cdot\sum_{r\in\{1,\cdots,m\}/c}\left(\frac{1}{j_r} + \mathbb{M}(n-2,j_r)\right)\right\}\\
&= n - \frac{1}{m} - \frac{1}{m}\cdot\sum_{r=1}^{m}\left(\frac{1}{j_r} + \mathbb{M}(n-2,j_r)\right) = W_\pi(n,t), \text{ and}\\
&\pi^*_{n-1}(t) \in d_1(t) = \pi_{n-1}(t).
\end{aligned} \tag{5.43}$$

The Value equation (5.12) under policy π satisfies the Bellman Equations. Hence, the policy π of marking a son is an optimal policy.

5.2 Budget two

With a budget of two, the problem is more complicated, as there are possibilities of marking any combination of sons and leaves.

Remark. There will be symmetrical trees in the usable state space U. We shall restrict the analysis to trees with increasing depth 1 subtree sizes, as any other combination of those depth 1 subtrees will result in the same value.

Lemma 5.5. *The set of usable states for budget 2 will contain all trees with unmarked leaves and not more than 2 marked depth-1 sons. i.e,* $\sum_{r=1}^{m} i_r \in \{0,1,2\}$.

Proof. We shall work with depth-1 subtrees of a given depth-2 tree.
Note that given a tree $t = (\mu, \{t_{i_1,j_1}, t_{i_2,j_2}, t_{i_3,j_3}, \cdots, t_{i_m,j_m}\})$, there will be m depth-

1 subtrees which are precisely $t_{i_1,j_1}, t_{i_2,j_2}, t_{i_3,j_3}, \cdots, t_{i_m,j_m}$. There is also the depth-1 subtree - root with m sons. But we do not bother about this subtree since it is not a subtree that will be "discovered" further. There are $(m+1)$ depth-1 subtrees in total. Since we start with unmarked trees at $n = 0$, and there is a budget of 2, no state after marking at $n = 1$ will have a depth 1 subtree with more than two marked sons. This implies $\sum_{r=1}^{m} i_r \leq 2$. Assume this statement holds true for any $n < N$. That is, all states after marking at $n < N$ do not have a depth 1 subtree with more than two marked sons.
To prove it for $n + 1$, let us proceed by contradiction.
Assume that there is a tree $t' = (\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\})$ with $\sum_{r=1}^{m} i_r > 2$, say that $\sum_{r=1}^{m} i_r = 3$. Without loss of generality, let $i_1, i_2, i_3 = 1$.
The tree t' is one of the discoveries of any of the m depth one subtrees (say t_b) of the tree at n after marking, and this t_b has three marked leaves. The root of t with its m sons is also a depth one subtree (contributing the +1 to the number of depth one subtrees), but we do not discover this subtree further! So t' is not one of the discoveries of the subtree $t_{\mu,m}$.
Precisely,

$$t' \in \mathcal{D}(t_b) = \mathcal{D}(\mu, \{1, 1, 1, 0, \cdots, 0)\}) \text{ there are three "1s" and } m - 3 \text{ "0s"}.$$

But t_b is a depth 1 subtree of a tree at n after marking, and according to our induction statement such a t_b does not exist. Hence the tree t' does not exist. $\square$

Let us calculate the Optimal policy for separate types of trees for each of the $1 \leq n \leq N$. To begin, let us label the actions in Table 5.2 :

Table 5.2: Table of Action Labels

Sl no.	Action Label	Action Interpretation
1	$a(d1, d1)$	Mark any two unmarked depth 1 nodes.
2	$a(d1, l_g)$	Mark any unmarked depth 1 node and leaf of tree t_{i_g,j_g} with $g \in \{1, \cdots, m\}$.
3	$a(l_g, l_h)$	Mark leaf of tree t_{i_g,j_g}, and leaf of tree t_{i_h,j_h} where $g, h \in \{1, \cdots, m\}$, $g \neq h$.
4	$a(l_g, l_g)$	Mark two leaves of tree t_{i_g,j_g} with $g \in \{1, \cdots, m\}$.

We now partition the state space to ease the analysis. The intuition for partitioning the state space in the way we have is the following.

Assume we prioritize marking the depth one nodes over marking leaves. Under such a marking, the trees that would still have unmarked depth one nodes, or all marked depth one nodes such that there is no more budget to mark leaves would be in one type.

A second type would be the trees where following the above marking would leave some budget that can be used to mark leaves. The second type is further subdivided into three types depending on the number of marked and umarked depth 1 sons.

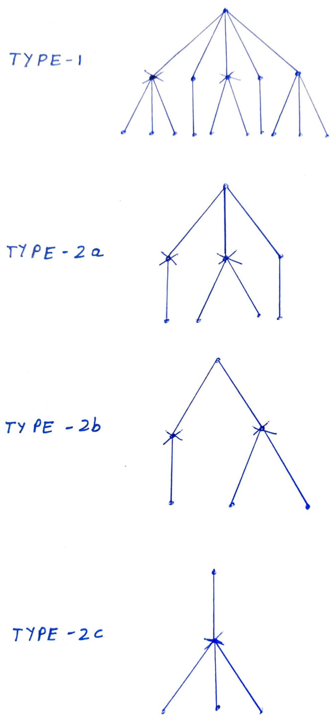

Figure 5.1: Types of depth 2 trees in the state space for budget 2

Figure 5.1 provides a visual depiction of the tree types. In the trees, the crosses represent markings.

5.2.1 Horizon $n = 1$

Type 1 - Trees with $m \geq \sum_{r=1}^{m} i_r + 2$:
In this case, it is optimal to use the budget to mark depth 1 nodes as that action has the minimum cost. The action set $A = \{a(d1, d1), a(d1, l_g), a(l_g, l_g), a(l_g, l_h)\}$ where $g, h \in \{1, \cdots, m\}$ and $g \neq h$. We find the minimum over A in order of the actions.

$$
\begin{aligned}
W^*(1, t) &= \min_{a \in A} \{c(t, a(t))\} \\
&= \min \left\{ \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r - 2), \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r - 1), \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r), \right. \\
&\left. \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r) \right\} \\
&= \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r - 2).
\end{aligned}
$$

Type 2a,2b,2c - Trees with $m < \sum_{r=1}^{m} i_r + 2$:
The action $a(d1, l_g)$ for Type $2a$ and $2b$ trees ensures that all depth 1 nodes are marked and hence results in a zero cost. Thus, $a(d1, l_g)$ leads to the minimum value for Type $2a$ and $2b$ as the other actions lead to a non-zero cost. For Type $2c$, the optimal action is, rather trivially, $a(l_1, l_1)$ since there is only one subtree attached to the root.

$$W^*(1, t) = 0 \text{ for all } t.$$

Summary of Optimal Values.

$$
W^*(1, t) = \begin{cases} \frac{1}{m}(m - \sum_{r=1}^{m} i_r - 2) & \text{for } m \geq \sum_{r=1}^{m} i_r + 2 \\ \\ 0 & \text{otherwise.} \end{cases} \tag{5.44}
$$

In Table 5.3 we have listed the corresponding optimal actions.

Table 5.3: Table for Optimal marking actions corresponding to $W^*(1,t)$

Tree Type	Optimal Marking Action
$m \geq \sum_{r=1}^{m} i_r + 2$	$a(d1, d1)$, Mark two depth one nodes.
$m - \sum_{r=1}^{m} i_r = 1$	$a(d1, l_{j_1})$, Mark unmarked depth one node and leaf of tree t_{i_1,j_1}.
$m - \sum_{r=1}^{m} i_r = 0$	$a(l_{j_g}, l_{j_h})$, where $g, h \in \{1, \cdots, m\}$, Mark any two leaves.

5.2.2 Horizon $n = 2$

Type 1 :

The action set is : $A = \{a(d1, d1), a(d1, l_g), a(l_g, l_h)\}$ where $g, h \in \{1, \cdots, m\}$ with the possibility that $g = h$.

Whenever we mark a leaf (or leaves) of a particular subtree, denote that subtree by t_{i_c,j_c} or by $(t_{i_{c1},j_{c1}}, t_{i_{c2},j_{c2}})$ if two leaves are marked.

We use the values for $n = 1$ from (5.44) in the Bellman Equations to obtain :

$$
\begin{aligned}
W^*(2,t) &= \min_{a \in A} \left\{ c(t, a(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t, a(t), t') \cdot W^*(1, t') \right\} \\
&= \min \left\{ \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r - 2) + \frac{1}{m} \cdot \sum_{r=1}^{m} \frac{(j_r - 2)^+}{j_r}, \right. && (5.45) \\
&\quad \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r - 1) + \frac{1}{m} \cdot \sum_{r \in \{1, \cdots, m\}/c} \frac{(j_r - 2)^+}{j_r} + \frac{1}{m} \cdot \frac{(j_c - 3)^+}{j_c}, && (5.46) \\
&\quad \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r) + \frac{1}{m} \cdot \sum_{r \in \{1, \cdots, m\}/\{c1, c2\}} \frac{(j_r - 2)^+}{j_r} \\
&\quad \left. + \frac{1}{m} \cdot \frac{(j_{c1} - 3)^+}{j_{c1}} + \frac{1}{m} \cdot \frac{(j_{c2} - 3)^+}{j_{c2}} \right\} && (5.47) \\
&= \frac{1}{m} \cdot (m - \sum_{r=1}^{m} i_r - 2) + \frac{1}{m} \cdot \sum_{r=1}^{m} \frac{(j_r - 2)^+}{j_r}.
\end{aligned}
$$

The minimum among the actions is $a(d1, d1)$. The expression in (5.45) is lesser than or equal to (5.46) because their difference is $1/m \cdot (1 - 1/j_c) > 0$ for $j_c \geq 3$, and 0 for $j_c < 3$. Similarly, the expression in (5.45) is lesser than or equal to (5.47) because their difference is $1/m \cdot (1 - 1/j_{c1}) + 1/m \cdot (1 - 1/j_{c2}) > 0$ for $j_{c1}, j_{c2} \geq 3$, and 0 for $j_{c1}, j_{c2} < 3$. If $j_{c1} < 3$ and $j_{c2} \geq 3$ then their difference is $1/m \cdot (1 - 1/j_{c2}) > 0$. This reasoning is also applied while calculating the minimum value among actions for Type $2a$ trees.

Type 2a :

The action set is $A = \{a(d1, l_g), a(l_g, l_h)\}$:

$$\begin{aligned}
W^*(2,t) &= \min_{a \in A} \left\{ c(t, a(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t, a(t), t') \cdot W^*(1, t') \right\} \\
&= \min\Big\{ \frac{1}{m} \cdot \sum_{r \in \{1,\cdots,m\}/c} \frac{(j_r - 2)^+}{j_r} + \frac{1}{m} \cdot \frac{(j_c - 3)^+}{j_c}, \\
&\quad \frac{1}{m} + \frac{1}{m} \cdot \sum_{r \in \{1,\cdots,m\}/\{c1,c2\}} \frac{(j_r - 2)^+}{j_r} + \frac{1}{m} \cdot \frac{(j_{c1} - 3)^+}{j_{c1}} \\
&\quad + \frac{1}{m} \cdot \frac{(j_{c2} - 3)^+}{j_{c2}} \Big\} \\
&= \frac{1}{m} \cdot \sum_{r \in \{1,\cdots,m\}/c} \frac{(j_r - 2)^+}{j_r} + \frac{1}{m} \cdot \frac{(j_c - 3)^+}{j_c}.
\end{aligned}$$

In continuation to the analysis of Type $2a$, we note that it is better to mark the leaf which has more than 2 siblings, and the minimum number of siblings. If all leaves have less than 2 siblings, then mark any leaf.

Let us look at why we have to mark the leaf which has more than 2 siblings, and the minimum number of siblings. Let $j_s = \min_{r \in \{1,\cdots,m\}} j_r$, and $j_l = \max_{r \in \{1,\cdots,m\}} j_r$ with $j_s, j_l \geq 3$. The difference between marking the leaf of subtree t_{i_s,j_s} and t_{i_l,j_l} is :

$$\left(\frac{j_s - 3}{j_s} + \sum_{r \in \{1,\cdots,m\}/\{s,l\}} \left(\frac{j_r - 2}{j_r} \right) + \frac{j_l - 2}{j_l} \right)$$

$$
\begin{aligned}
&- \left(\frac{j_s-2}{j_s} + \sum_{r\in\{1,\cdots,m\}/\{s,l\}} \left(\frac{j_r-2}{j_r}\right) + \frac{j_l-3}{j_l} \right) \\
&= \frac{-1}{j_s} + \frac{-1}{j_l} \\
&< 0 \text{ since } j_s < j_l.
\end{aligned}
$$

This implies that it is better to mark the leaf of subtree t_{i_s,j_s}.
Hence, mark j_c such that

$$
j_c = \max\{\min_r\{j_r \geq 3\}, \max_r\{j_r \leq 2\}\} \text{ where } r \in \{1, \cdots, m\}. \tag{5.48}
$$

Type 2b :
The relevant actions here are marking the leaves since the sons are marked already. The particular leaf to mark is what we aim to find.
Let $A = \{$ (First action) $a(l_{j_1}, l_{j_2})$, (Second action) $a(l_{j_1}, l_{j_1})$, (Third action) $a(l_{j_2}, l_{j_2})$ $\}$. Let us solve the Bellman equation for the actions in this order.

$$
\begin{aligned}
W^*(2,t) &= \min_{a\in A} \left\{ c(t,a(t)) + \sum_{t'\in \mathcal{SD}(t)} P(t,a(t),t') \cdot W^*(1,t') \right\} \\
&= \min \left\{ \frac{1}{m} \cdot \frac{(j_1-3)^+}{j_1} + \frac{1}{m} \cdot \frac{(j_2-3)^+}{j_2}, \right. \\
&\left. \frac{1}{m} \cdot \frac{(j_1-4)^+}{j_1} + \frac{1}{m} \cdot \frac{(j_2-2)^+}{j_2}, \frac{1}{m} \cdot \frac{(j_1-2)^+}{j_1} + \frac{1}{m} \cdot \frac{(j_2-4)^+}{j_2} \right\}
\end{aligned}
$$

$$
= \begin{cases}
\frac{1}{m} \cdot \frac{(j_1-3)^+}{j_1} + \frac{1}{m} \cdot \frac{(j_2-3)^+}{j_2} & \text{if } j_1 \text{ or } j_2 = 3 \\
\frac{1}{m} \cdot \frac{(j_1-4)^+}{j_1} + \frac{1}{m} \cdot \frac{(j_2-2)^+}{j_2} & \text{if } j_1, j_2 > 3 \\
\frac{1}{m} \cdot \frac{(j_2-4)^+}{j_2} & \text{if } j_1 < 3 < j_2 \\
0 & \text{if } j_1, j_2 < 3
\end{cases}
$$

The optimal decision for Type $2b$ trees can be described as follows :

- If both subtrees have more than three leaves, then mark the leaves of t_{i_1,j_1} i.e., the smallest subtree.

- If either of the subtrees have three leaves, mark one of those leaves and the leaf of the other subtree.

- If $j_1 < 3$, then it is better to mark two leaves of t_{i_2,j_2}.

- If both subtrees have less than three leaves, we can mark any pair of leaves since the cost is 0 for any combination of marking leaves.

Type 2c : Trivially, mark any two leaves.

$$W^*(2,t) = \frac{1}{m} \cdot \frac{(j_1-4)^+}{j_1} = \frac{(j_1-4)^+}{j_1}. \tag{5.49}$$

Table 5.4 lists the optimal actions for each tree type which is followed by the summary of optimal value functions for all tree types.

Table 5.4: Table for Optimal marking actions corresponding to $W^*(2,t)$

Tree Type	Optimal Marking Action
$m \geq \sum_{r=1}^{m} i_r + 2$	$a(d1, d1)$, Mark any two unmarked depth 1 nodes.
$m - \sum_{r=1}^{m} i_r = 1$	$a(d1, l_{j_c})$, Mark unmarked depth 1 node and leaf of tree t_{i_c,j_c} with c defined in (5.48)
$m = 2, \sum_{r=1}^{m} i_r = 2, j_1$ or $j_2 = 3$	$a(l_{j_1}, l_{j_2})$, Mark one leaf each of subtrees t_{i_1,j_1}, t_{i_2,j_2} .
$m = 2, \sum_{r=1}^{m} i_r = 2, j_1, j_2 > 3$	$a(l_{j_1}, l_{j_1})$, Mark two leaves of subtree t_{i_1,j_1}.
$m = 2, \sum_{r=1}^{m} i_r = 2, j_1, j_2 < 3$	Mark any two leaves.
$m = 2, \sum_{r=1}^{m} i_r = 2, j_1 < 3 < j_2$	$a(l_{j_2}, l_{j_2})$, Mark two leaves of subtree t_{i_2,j_2}.
$m = 1, \sum_{r=1}^{m} i_r = 1$	$a(l_{j_1}, l_{j_1})$ Mark any two leaves of the sole subtree.

Summary of Optimal Value functions at $n=2$ for each tree type:

$$W^*(2,t) = \begin{cases} \frac{1}{m}(m - \sum_{r=1}^{m} i_r - 2) & \\ +\frac{1}{m} \cdot \sum_{r=1}^{m} \frac{1}{j_r}(j_r - 2)^+ & \text{for } m \geq \sum_{r=1}^{m} i_r + 2 \\ \frac{1}{m} \sum_{r \in \{1,\cdots,m\}/c} \frac{(j_r-2)^+}{j_r} & \text{for } m - \sum_{r=1}^{m} i_r = 1 \text{ where,} \\ +\frac{1}{m} \cdot \frac{(j_c - 3)^+}{j_c} & j_c \text{ as in (5.48)} \\ \frac{1}{m} \cdot \frac{(j_1-3)^+}{j_1} + \frac{1}{m} \cdot \frac{(j_2-3)^+}{j_2} & \text{for } m = 2, \sum_{r=1}^{m} i_r = 2, j_1 \text{ or } j_2 = 3 \\ \frac{1}{m} \cdot \frac{(j_1-4)^+}{j_1} + \frac{1}{m} \cdot \frac{(j_2-2)^+}{j_2} & \text{for } m = 2, \sum_{r=1}^{m} i_r = 2, j_1, j_2 > 3 \\ \frac{1}{m} \cdot \frac{(j_2-4)^+}{j_2} & \text{for } m = 2, \sum_{r=1}^{m} i_r = 2, j_1 < 3 < j_2 \\ 0 & \text{for } m = 2, \sum_{r=1}^{m} i_r = 2, j_1, j_2 < 3 \\ \frac{1}{m} \cdot \frac{(j_1-4)^+}{j_1} & \text{for } m = 1, \sum_{r=1}^{m} i_r = 1. \end{cases} \tag{5.50}$$

5.2.3 Horizon $n = 3$

Preliminary Computations

We shall define and then compute the closed form solutions (upto $\mathbb{H}_p$) of the following terms. We refer to these terms frequently in this section.

$$
\begin{aligned}
Tr(2:0,0) &= \sum_{f_1=1}^{p}\sum_{f_2=1}^{p} W^*(2,(i_c,\{t_{0,f_1},t_{0,f_2}\}))\\
Tr(2:1,0) &= \sum_{f_1=1}^{p}\sum_{f_2=1}^{p} W^*(2,(i_c,\{t_{1,f_1},t_{0,f_2}\}))\\
Tr(2:1,1) &= \sum_{f_1=1}^{p}\sum_{f_2=1}^{p} W^*(2,(i_c,\{t_{1,f_1},t_{1,f_2}\}))\\
Tr(3:1,1,0) &= \sum_{f_1=1}^{p}\sum_{f_2=1}^{p}\sum_{f_3=1}^{p} W^*(2,(i_c,\{t_{1,f_1},t_{1,f_2},t_{0,f_3}\})).
\end{aligned}
\tag{5.51}
$$

One should understand $Tr(2:1,0)$ as the accumulated values of all trees with two sons, one of which is marked, and the other unmarked. We adopt similar interpretation for the other terms.

$$
Tr(2:0,0) = p * \mathbb{H}_{p2} = p(p - 2\mathbb{H}_p + 1). \tag{5.52}
$$

Let us write the equation line by line, where each line is for a particular value of f_1 as indicated in the square brackets.

$$
\begin{aligned}
&2 \cdot Tr(2:1,0)\\
&= \sum_{f_2=1}^{p} \frac{(f_2-3)^+}{f_2}\ [f_1 = 1] && (5.53)\\
&+ \sum_{f_2=1}^{p} \frac{(f_2-3)^+}{f_2}\ [f_1 = 2] && (5.54)\\
&+ \sum_{f_2=1}^{p} \left(\frac{(f_1-3)^+}{f_1} + \frac{(f_2-2)^+}{f_2}\right)\ [f_1 = 3] && (5.55)\\
&+ \sum_{f_2=1}^{2} \frac{(f_1-3)^+}{f_1} + (\sum_{f_2=3}^{4} \frac{(f_1-2)^+}{f_1} + \frac{(f_2-3)^+}{f_2})\\
&+ \sum_{f_2=5}^{p} \left(\frac{(f_1-3)^+}{f_1} + \frac{(f_2-2)^+}{f_2}\right)\ [f_1 = 4] && (5.56)\\
&+ \sum_{f_2=1}^{2} \frac{(f_1-3)^+}{f_1} + \sum_{f_2=3}^{5} \left(\frac{(f_1-2)^+}{f_1} + \frac{(f_2-3)^+}{f_2}\right)
\end{aligned}
$$

$$+\sum_{f_2=6}^{p}\left(\frac{(f_1-3)^+}{f_1}+\frac{(f_2-2)^+}{f_2}\right)\ [f_1=5]$$

$\ldots$

$$+\sum_{f_2=1}^{2}\frac{(f_1-3)^+}{f_1}+\sum_{f_2=3}^{p}\left(\frac{(f_1-2)^+}{f_1}+\frac{(f_2-3)^+}{f_2}\right)[f_1=p]. \tag{5.57}$$

The summation is split into three cases beginning from $f_1 = 4$ in (5.56). This is so because as f_2 varies from 1 to p, we refer to (5.50) for the optimal value function. The optimal value is different for each of the cases : $f_2 = 1, 2,\ f_2 = 3,\ f_2 \geq f_1$, and $4 \leq f_2 \leq f_1$ since the optimal actions for these cases are different as noted in Table 5.4.

Let us add (5.53), (5.54), and (5.55), and then group the terms from the $f_1 = 4$ case based on the summations.

$$2\cdot Tr(2:1,0)=\mathbb{H}_{p3}+\mathbb{H}_{p3}+\mathbb{H}_{p2}+\sum_{f_1=4}^{p}2\frac{(f_1-3)^+}{f_1}$$
$$+\sum_{f_1=4}^{p}\sum_{f_2=3}^{f_1}\left(\frac{(f_1-2)^+}{f_1}+\frac{(f_2-3)^+}{f_2}\right)+\sum_{f_1=4}^{p-1}\sum_{f_2=f_1+1}^{p}\left(\frac{(f_1-3)^+}{f_1}+\frac{(f_2-2)^+}{f_2}\right)$$

$$=4\mathbb{H}_{p3}+\mathbb{H}_{p2}+\sum_{f_1=4}^{p}\sum_{f_2=3}^{f_1}\left(\frac{(f_1-2)^+}{f_1}+\frac{(f_2-3)^+}{f_2}\right)$$
$$+\sum_{f_1=4}^{p-1}\sum_{f_2=f_1+1}^{p}\left(\frac{(f_1-3)^+}{f_1}+\frac{(f_2-2)^+}{f_2}\right). \tag{5.58}$$

The expression in (5.58) is solved on Mathematica and the output is :

$$Tr(2:1,0)=p^2+\frac{7p}{2}+\frac{7}{4}-\frac{5\mathbb{H}_p}{2}-3p\mathbb{H}_p. \tag{5.59}$$

We compute $Tr(2:1,1), Tr(3:1,1,0)$ in a method similar to that of $Tr(2:1,0)$. The explicit computations for $Tr(2:1,1)$ is in Appendix A.3, and for $Tr(3:1,1,0)$ it is in Appendix A.4.

To summarise the results of the preliminary computations we have :

$$Tr(2:0,0) = p(p - 2\mathbb{H}_p + 1) \tag{5.60}$$

$$Tr(2:1,0) = p^2 + \frac{7p}{2} + \frac{7}{4} - \frac{1}{2}(5 + 6p)\mathbb{H}_p \tag{5.61}$$

$$Tr(2:1,1) = p^2 + \frac{19p}{3} + 5 - (4p + 6)\mathbb{H}_p \tag{5.62}$$

$$Tr(3:1,1,0) = \frac{5}{2} + 8p + 4p^2 + p^3 - \frac{1}{3}(19 + 15p + 9p^2)\mathbb{H}_p. \tag{5.63}$$

Analysis

Type 1 :

To begin with, let us calculate the values of Type 1 trees under different actions and then compare them. The value function of Type 1 trees under action $a(d1, d1)$ is:

$$\begin{aligned}
W_{a(d1,d1)}(3,t) &= c(t, a(d1,d1)(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t, a(d1,d1)(t), t') \cdot W^*(2,t') \\
&= c(t, a(d1,d1)(t)) + \sum_{r=1}^{m} \frac{1}{m \cdot p^{j_r}} \sum_{t' \in \mathcal{D}(t_{i_r,j_r})} W^*(2,t') \\
&= c(t, a(d1,d1)(t)) \\
&+ \sum_{r=1}^{m} \left(\frac{1}{m \cdot p^{j_r}} \sum_{1 \le f_1, \cdots, f_{j_r} \le p} W^*(2, (\mu, \{t_{0,f_1}, \cdots, t_{0,f_{j_r}}\})) \right).
\end{aligned} \tag{5.64}$$

Let us split the term $W^*(2, (\mu, \{t_{0,f_1}, t_{0,f_2}, \cdots, t_{0,f_{j_r}}\})$ based on the values j_r could take. We split this by referring to the summary of optimal decisions at $n = 2$ in (5.50). The optimal value functions for $j_r \ge 2$ and $j_r = 1$ are different since the optimal actions differ as noted in Table 5.4.

$$
\begin{aligned}
W_{a(d1,d1)}(3,t) &= c(t, a(d1,d1)(t)) \\
&+ \sum_{r=1}^{m} \mathbb{1}_{\{j_r \geq 2\}} \frac{1}{m \cdot p^{j_r}} \sum_{1 \leq f_1, \cdots, f_{j_r} \leq p} \left(\frac{j_r - 2}{j_r} + \frac{1}{j_r} \sum_{q=1}^{j_r} \frac{(f_q - 2)^+}{f_q} \right) \\
&+ \sum_{r=1}^{m} \mathbb{1}_{\{j_r = 1\}} \frac{1}{m \cdot p} \sum_{1 \leq f_1 \leq p} \frac{(f_1 - 3)^+}{f_1} \\
&= c(t, a(d1,d1)(t)) \\
&+ \frac{1}{m} \cdot \sum_{r=1}^{m} \left(\mathbb{1}_{\{j_r \geq 2\}} \left(\frac{(j_r - 2)^+}{j_r} + \frac{1}{p} \mathbb{H}_{p2} \right) + \mathbb{1}_{\{j_r = 1\}} \left(\frac{1}{p} \mathbb{H}_{p3} \right) \right),
\end{aligned}
\tag{5.65}
$$

where $\mathbb{H}_{pk}$ is as defined in (4.32). The explicit computation by which we obtain the term $p^{-1}\mathbb{H}_{p2}$ when $j_r \geq 2$ is in Appendix A.5.

Under the action $a(d1, l_c)$ where $c \in \{j_1, \cdots, j_m\}$, the value is :

$$
\begin{aligned}
W_{a(d1,l_c)}(3,t) &= c(t, a(d1, l_{j_c})(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t, a(d1, l_{j_c})(t), t') \cdot W^*(2, t') \\
&= c(t, a(d1, l_{j_c})(t)) + \sum_{r=1}^{m} \frac{1}{m \cdot p^{j_r}} \sum_{t' \in \mathcal{D}(t_{i_r, j_r})} W^*(2, t').
\end{aligned}
\tag{5.66}
$$

Split the summation into two parts (with sub-parts). The first part runs over $\{1, \cdots, m\}/c$ which is the discovery of sub-trees with no marked leaf, and the other part is discovery of the sub-tree t_{i_c, j_c} which has one marked leaf. The sub-parts in each part are for different values of j_r and j_c. Therefore, we have :

$$
\begin{aligned}
W_{a(d1,l_c)}(3,t) = \frac{m - \sum_{r=1}^{m} i_r - 1}{m} &+ \frac{1}{m} \sum_{r \in \{1, \cdots, m\}/c} \left(\mathbb{1}_{\{j_r \geq 2\}} \left(\frac{(j_r - 2)^+}{j_r} + \frac{1}{p} \mathbb{H}_{p2} \right) \right. \\
&\left. + \mathbb{1}_{\{j_r = 1\}} \left(\frac{1}{p} \mathbb{H}_{p3} \right) \right) + \frac{1}{m} \cdot \left(\mathbb{1}_{\{j_c \geq 3\}} \left(\frac{(j_c - 3)^+}{j_c} + \frac{1}{p} \mathbb{H}_{p2} \right) \right) \\
&+ \frac{1}{mp^2} \cdot \mathbb{1}_{\{j_c = 2\}} \sum_{f_1=1}^{p} \sum_{f_2=1}^{p} W^*(2, (i_c, \{t_{1,f_1}, t_{0,f_2}\})) \\
&+ \mathbb{1}_{\{j_c = 1\}} \frac{1}{mp} \cdot \sum_{f_1=1}^{p} \frac{(f_1 - 4)^+}{f_1}.
\end{aligned}
\tag{5.67}
$$

To ease the readability of the value equation, we shall introduce the following notation.

$$lun(j) = \mathbb{1}_{\{j\geq 2\}}\left(\frac{(j-2)^+}{j} + \frac{1}{p}\mathbb{H}_{p2}\right) + \mathbb{1}_{\{j=1\}}\left(\frac{1}{p}\mathbb{H}_{p3}\right) \tag{5.68}$$

$$lm1(j) = \left(\frac{(j-3)^+}{j} + \frac{1}{p}\mathbb{H}_{p2}\right) \tag{5.69}$$

$$lm2(j) = \left(\frac{(j-4)^+}{j} + \frac{1}{p}\mathbb{H}_{p2}\right) \tag{5.70}$$

The notation $lun(j)$ should be understood as costs of "leaf-unmarked" which is the costs to go of the discoveries of trees $t_{i,j}$ with unmarked leaves. Similarly $lm1(j)$ is the costs to go of the discoveries of trees $t_{i,j}$ with one marked leaf. The same interpretation applies for $lm2(j)$.

With this notation, (5.65) and (5.67) can be written as:

$$W_{a(d1,d1)}(3,t) = \frac{m - \sum_{r=1}^{m} i_r - 2}{m} + \frac{1}{m}\sum_{r=1}^{m} lun(j_r) \tag{5.65}$$

$$\begin{aligned} W_{a(d1,l_c)}(3,t) = {} & \frac{m - \sum_{r=1}^{m} i_r - 1}{m} + \frac{1}{m}\sum_{r\in\{1,\cdots,m\}/c} lun(j_r) + \frac{1}{m}\,\mathbb{1}_{\{j_c\geq 3\}}\, lm1(j_c) \\ & + \frac{1}{mp^2}\mathbb{1}_{\{j_c=2\}}\, Tr(2:1,0) + \mathbb{1}_{\{j_c=1\}}\frac{1}{mp}\cdot\sum_{f_1=1}^{p}\frac{(f_1-4)^+}{f_1}. \end{aligned} \tag{5.67}$$

Now, let us compare the values of the actions $a(d1,d1)$ and $a(d1,l_c)$ for each of the cases $j_c \geq 3$, $j_c = 2$ and $j_c = 1$.

For $j_c \geq 3$. Difference between (5.65) and (5.67) is :

$$\begin{aligned} W_{a(d1,l_c)}(3,t) - W_{a(d1,d1)}(3,t) &= \frac{1}{m} + \frac{1}{m}\cdot\left(\left(\frac{(j_c-3)^+}{j_c} + \frac{1}{p}\mathbb{H}_{p2}\right)\right. \\ &\quad \left. - \left(\frac{(j_c-2)^+}{j_c} + \frac{1}{p}\mathbb{H}_{p2}\right)\right) \\ &= \frac{1}{m}\cdot\left(1 - \frac{1}{j_c}\right) \\ &> 0. \end{aligned} \tag{5.71}$$

For $j_c = 1$, we have :

$$W_{a(d1,l_c)}(3,t) - W_{a(d1,d1)}(3,t) = \frac{1}{m} + \frac{1}{mp}\cdot(\mathbb{H}_{p4} - \mathbb{H}_{p3})$$

$$= \frac{1}{m} \cdot \left(1 + \frac{(\mathbb{H}_p - \mathbb{H}_3)^+}{p}\right) \tag{5.72}$$
$$\geq 0.$$

The term in (5.72) is obviously non-negative for $p \geq 3$ since $\mathbb{H}_p - \mathbb{H}_3 \geq 0$, and for $p = 1, 2$, $\mathbb{H}_p - \mathbb{H}_3 \geq -1$.

For $j_c = 2$, we have :

$$W_{a(d1,l_c)}(3,t) - W_{a(d1,d1)}(3,t) = \frac{1}{m} + \frac{1}{mp^2} Tr(2:1,0) - \frac{1}{mp^2} Tr(2:0,0). \tag{5.73}$$

Thus, if we can prove that (5.73) is Strictly Positive, $a(d1, d1)$ dominates $a(d1, l_c)$.

(5.73) can be written as,

$$\frac{1}{m} \left\{ \frac{4p^3 + 14p^2 - 8p^2\mathbb{H}_p - 10p\mathbb{H}_p + 3p}{4p^3} \right\}. \tag{5.74}$$

To check that the action $a(d1, d1)$ is better, we only need to show that the term within braces is non-negative. i.e., we need to show that,

$$4p^2(p - 2\mathbb{H}_p) + 2p(7p - 10\mathbb{H}_p) + 3p \geq 0 \text{ for all } p. \tag{5.75}$$

First, let us check that (5.75) holds for all $p \geq 5$. It can be easily seen that for $p \geq 5$, the following holds true :

$$p > 2\mathbb{H}_p > \frac{10}{7}\mathbb{H}_p.$$

For $p = 1, 2, 3, 4$, the left hand side of (5.75) evaluates to 3, 16, 56, and 142 respectively.
Hence (5.73) is strictly positive for all $p \geq 1$.

We have to assert that the action $a(l_{c1}, l_{c2})$ with the possibility that $c1 = c2 = c$ (marking two leaves) is not optimal as well before concluding that $a(d1, d1)$ is the optimal action. To do this we shall show that the $a(l_{c1}, l_{c2})$ is actually worse than the action $a(d1, l_c)$.
The value function under action $a(d1, l_c)$ is in (5.67).

For action $a(l_c, l_c)$, where we mark two leaves of the same subtree, the value function is :

$$
\begin{aligned}
W_{a(l_c,l_c)}(3,t) = {} & \frac{m - \sum_{r=1}^{m} i_r}{m} + \frac{1}{m} \sum_{r\in\{1,\cdots,m\}/\{c\}} lun(j_r) \\
& + \frac{1}{m} \cdot \mathbb{1}_{\{j_c \geq 4\}} lm2(j_c) + \frac{1}{mp^3} \cdot (\mathbb{1}_{\{j_c=3\}}) Tr(3:1,1,0) \\
& + \frac{1}{mp^2} \cdot (\mathbb{1}_{\{j_c=2\}}) Tr(2:1,1) + \frac{1}{mp} \cdot (\mathbb{1}_{\{j_c=1\}}) \sum_{f_1=1}^{p} \frac{(f_1-4)^+}{f_1}.
\end{aligned}
\tag{5.76}
$$

For marking leaves of different subtrees the value function is,

$$
\begin{aligned}
W_{a(l_{c1},l_{c2})}(3,t) = {} & \frac{m - \sum_{r=1}^{m} i_r}{m} + \frac{1}{m} \sum_{r\in\{1,\cdots,m\}/\{c1,c2\}} lun(j_r) \\
& + \frac{1}{m} \cdot \left(\mathbb{1}_{\{j_{c1} \geq 3\}} lm1(j_{c1})\right) + \frac{1}{m} \cdot \left(\mathbb{1}_{\{j_{c2} \geq 3\}} lm1(j_{c2})\right) \\
& + \frac{1}{mp^2} \cdot (\mathbb{1}_{\{j_{c1}=2\}} + \mathbb{1}_{\{j_{c2}=2\}}) Tr(2:1,0) \\
& + \frac{1}{mp} \cdot (\mathbb{1}_{\{j_{c1}=1\}} + \mathbb{1}_{\{j_{c2}=1\}}) \sum_{f_1=1}^{p} \frac{(f_1-4)^+}{f_1}.
\end{aligned}
\tag{5.77}
$$

Before comparing the values under the three actions, we assume without loss of generality, that one of the leaves marked is common in all three actions. That is, $c = c1$ in $a(d1, l_c), a(l_c, l_c)$ and $a(l_{c1}, l_{c2})$. Among the actions $a(l_{j_c}, l_{j_c})$ and $a(l_{j_{c1}}, l_{j_{c2}})$, there would be some particular $a(l_{j_c}, l_{j_c})$ that would dominate over all other $a(l_{j_g}, l_{j_g})$ for $g \neq c$, and some particular $a(l_{j_{c1}}, l_{j_{c2}})$ would dominate over all other $a(l_{j_g}, l_{j_h})$ for $(g, h) \neq (c1, c2)$. Let us assume that when we pick the best action among the $a(l_{j_{c1}}, l_{j_{c1}})$ and $a(l_{j_{c1}}, l_{j_{c2}})$ actions, we compare these "best" actions with $a(d1, l_c)$ with a common leaf. We first establish that $a(d1, l_c)$ is better than $a(l_c, l_c)$ for any l_c. We also prove that $a(d1, l_{c1})$ is better than $a(l_{c1}, l_{c2})$ for all $c1 \neq c2$. Hence, we would obtain that $a(d1, l_{c1})$ is better than $a(l_{c1}, l_{c2})$ for any $c1, c2$. Note that even if $a(d1, l_c)$ is better than the best $a(l_{j_c}, l_{j_c})$ and $a(l_{j_{c1}}, l_{j_{c2}})$, $a(d1, l_c)$ is not necessarily the optimal action. There could be an action $a(d1, l_s)$ where $s \neq c$ which is optimal. However, in order to eliminate the actions $a(l_{j_{c1}}, l_{j_{c1}})$

and $a(l_{j_{c1}}, l_{j_{c2}})$, it is sufficient to show that $a(d1, l_c)$ is better than the best $a(l_{j_c}, l_{j_c})$ and $a(l_{j_{c1}}, l_{j_{c2}})$.

$$
\begin{aligned}
&W_{a(l_{c1},l_{c1})}(3,t) - W_{a(d1,l_{c1})}(3,t) \\
&= \begin{cases}
\dfrac{1}{m} - \dfrac{1}{m} \cdot \dfrac{1}{j_{c1}} & \text{for } j_{c1} \geq 4 \\[2ex]
\dfrac{1}{m} + \dfrac{1}{mp^3} \cdot Tr(3:1,1,0) - \dfrac{1}{mp}\left(p - 2\mathbb{H}_p + 1\right) & \text{for } j_{c1} = 3 \\[2ex]
\dfrac{1}{m} + \dfrac{1}{mp^2} \cdot Tr(2:1,1) - \dfrac{1}{mp^2} \cdot Tr(2:1,0) & \text{for } j_{c1} = 2 \\[2ex]
\dfrac{1}{m} & \text{for } j_{c1} = 1.
\end{cases}
\end{aligned} \tag{5.78}
$$

For the case $j_{c1} \geq 4$, the difference is obviously positive.

For $j_{c1} = 3$, we have, using (5.60) and (5.63), the following simplified difference which is positive:

$$
\frac{1}{m}\left(1 + \frac{1}{p^3}\left(\frac{5}{2} + 8p + 3p^2 - \mathbb{H}_p\left(\frac{19}{3} + 5p + p^2\right)\right)\right) > 0. \tag{5.79}
$$

For the $j_{c1} = 1$ case the action $a(l_{c1}, l_{c1})$ may not make sense since there is only one leaf. We assume that the action marks the sole leaf, and the other budget is unused. In such a case, the difference is positive rather trivially.

For $j_{c1} = 2$, the difference between values under actions $a(l_{c1}, l_{c1})$ and $a(d1, l_{c1})$ simplifies to:

$$
\begin{aligned}
&\frac{1}{m}\left(1 + \frac{1}{p^2}\left(Tr(2:1,1) - Tr(2:1,0)\right)\right) \\
&= \frac{1}{m \cdot 12p^2}\left(12p^2 - 12p\mathbb{H}_p + 34p - 42\mathbb{H}_p + 39\right).
\end{aligned} \tag{5.80}
$$

Again using that $p > 2\mathbb{H}_p > 42/34 \cdot \mathbb{H}_p > \mathbb{H}_p$, the expression in (5.80) is strictly positive. Hence,

$$
W_{a(l_{c1},l_{c1})}(3,t) > W_{a(d1,l_{c1})}(3,t) \text{ for all } c1 \in \{j_1, \cdots, j_m\}, \text{ and } t \in U.
$$

Similarly, we compare the actions $a(l_{c1}, l_{c2})$ and $a(d1, l_{c1})$:

$$W_{a(l_{c1},l_{c2})}(3,t) - W_{a(d1,l_{c1})}(3,t)$$
$$= \begin{cases} \dfrac{1}{m} - \dfrac{1}{m} \cdot \dfrac{1}{j_{c2}} & \text{for } j_{c2} \geq 3 \\[2ex] \dfrac{1}{m} + \dfrac{1}{mp^2} \cdot Tr(2:1,0) - \dfrac{1}{mp^2} \cdot Tr(2:0,0) & \text{for } j_{c2} = 2 \\[2ex] \dfrac{1}{m} & \text{for } j_{c2} = 1. \end{cases} \tag{5.81}$$

For $j_{c2} = 2$, we have the reasoning from (5.73) that the difference is non-negative. Hence,

$$W_{a(l_{c1},l_{c2})}(3,t) > W_{a(d1,l_{c1})}(3,t) \text{ for all } c1, c2 \in \{j_1, \cdots, j_m\}, \text{ and } t \in U.$$

Lemma 5.6. *The optimal action for Type 1 trees at $n = 3$ is $a(d1, d1)$, which is marking two unmarked depth 1 nodes.*

Type 2 :

In case of marking leaves, we shall initially consider that any leaf(s) is(are) being marked. Later on, we shall find which leaf(s) to mark.

Type 2a :

The set of relevant actions are : $A = \{a(d1, l_g), a(l_g, l_h)\}$ where $g, h \in \{j_1, \cdots, j_m\}$ with the possibility that $g = h$.

Under action $a(d1, l_c)$:

$$\begin{aligned} W_{a(d1,l_c)}(3,t) &= c(t, a(d1,l_c)(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t, a(d1,l_c)(t), t') \cdot W^*(2,t') \\ &= c(t, a(d1,l_c)(t)) + \sum_{r=1}^{m} \frac{1}{m \cdot p^{j_r}} \sum_{t' \in \mathcal{D}(t_{i_r,j_r})} W^*(2,t') \\ &= c(t, a(d1,l_c)(t)) + \\ &\quad \sum_{r \in \{1,\cdots,m\}/c} \frac{1}{mp^{j_r}} \sum_{1 \leq f_1,\cdots,f_{j_r} \leq p} W^*(2, (\mu, \{t_{0,f_1}, \cdots, t_{0,f_{j_r}}\})) \end{aligned}$$

$$+ \frac{1}{mp^{j_c}} \sum_{1 \le f_1, \cdots, f_{j_c} \le p} W^*(2, (\mu, \{t_{1,f_1}, \cdots, t_{0,f_{j_c}}\}))$$

Using indicators for different j_r and j_c:

$$\begin{aligned} W_{a(d1,l_c)}(3,t) &= \frac{1}{m} \sum_{r \in \{1,\cdots,m\}/c} lun(j_r) + \frac{1}{m} \cdot \left(\mathbb{1}_{\{j_c \ge 3\}} lm1(j_c)\right) \\ &+ \frac{1}{mp^2} \cdot \mathbb{1}_{\{j_c=2\}} Tr(2:1,0) + \mathbb{1}_{\{j_c=1\}} \frac{1}{mp} \cdot \sum_{f_1=1}^{p} \frac{(f_1-4)^+}{f_1}. \end{aligned} \tag{5.82}$$

Under action $a(l_{c1}, l_{c2})$ (mark leaves of different subtrees):

$$\begin{aligned} W_{a(l_{c1},l_{c2})}(3,t) &= \frac{1}{m} \\ &+ \sum_{r \in \{1,\cdots,m\}/\{c1,c2\}} \frac{1}{mp^{j_r}} \sum_{1 \le f_1,\cdots,f_{j_r} \le p} W^*(2, (\mu, \{t_{0,f_1}, \cdots, t_{0,f_{j_r}}\})) \\ &+ \frac{1}{mp^{j_{c1}}} \sum_{1 \le f_1,\cdots,f_{j_{c1}} \le p} W^*(2, (\mu, \{t_{1,f_1}, \cdots, t_{0,f_{j_{c1}}}\})) \\ &+ \frac{1}{mp^{j_{c2}}} \sum_{1 \le f_1,\cdots,f_{j_{c2}} \le p} W^*(2, (\mu, \{t_{1,f_1}, \cdots, t_{0,f_{j_{c2}}}\})). \end{aligned} \tag{5.83}$$

Splitting the cases for different j_r and j_{c1}, j_{c2} using indicators, we get,

$$\begin{aligned} W_{a(l_{c1},l_{c2})}(3,t) &= \frac{1}{m} + \frac{1}{m} \sum_{r \in \{1,\cdots,m\}/\{c1,c2\}} lun(j_r) \\ &+ \frac{1}{m} \cdot \left(\mathbb{1}_{\{j_{c1} \ge 3\}} lm1(j_{c1})\right) + \frac{1}{m} \cdot \left(\mathbb{1}_{\{j_{c2} \ge 3\}} lm1(j_{c2})\right) \\ &+ \frac{1}{mp^2} \cdot (\mathbb{1}_{\{j_{c1}=2\}} + \mathbb{1}_{\{j_{c2}=2\}}) Tr(2:1,0) \\ &+ \frac{1}{mp} \cdot (\mathbb{1}_{\{j_{c1}=1\}} + \mathbb{1}_{\{j_{c2}=1\}}) \sum_{f_1=1}^{p} \frac{(f_1-4)^+}{f_1}. \end{aligned} \tag{5.84}$$

A similar calculation for the action of marking two leaves of same subtree $a(l_c, l_c)$ (without loss of generality, assume first two leaves are marked):

$$
\begin{aligned}
W_{a(l_c,l_c)}(3,t) = \frac{1}{m} + \frac{1}{m} &\sum_{r\in\{1,\cdots,m\}/\{c\}} lun(j_r) \\
&+ \frac{1}{m}\cdot\left(\mathbb{1}_{\{j_c\geq 4\}} lm2(j_c)\right) + \frac{1}{mp^3}\cdot(\mathbb{1}_{\{j_c=3\}})Tr(3:1,1,0) \\
&+ \frac{1}{mp^2}\cdot(\mathbb{1}_{\{j_c=2\}})Tr(2:1,1) + \frac{1}{mp}\cdot(\mathbb{1}_{\{j_c=1\}})\sum_{f_1=1}^{p}\frac{(f_1-4)^+}{f_1}.
\end{aligned}
\tag{5.85}
$$

We assume without loss of generality that one of the leaves marked is common in all three actions. That is, $c = c1$ in $a(d1, l_c), a(l_c, l_c)$ and $a(l_{c1}, l_{c2})$.
The difference between (5.84) and (5.82) is exactly (5.81).
The difference between (5.85) and (5.82) is exactly (5.78).

The conclusion of the formal analysis of actions $a(d1, l_{c1})$, $a(l_{c1}, l_{c1})$ and $a(l_{c1}, l_{c2})$ for Type $2a$ trees has been stated in the following Lemma 5.7.

Lemma 5.7. *For type 2a trees, the optimal action is $a(d1, l_c)$, that is marking the unmarked depth 1 node and any leaf is better than marking any combination of leaves.*

Let us find the specific leaf to mark according to Lemma 5.7. First consider trees where $j_r \geq 3$ for at least one $r \in \{1, \cdots, m\}$.
Among all the $j_r \geq 3$, it is better to mark the leaf of tree t_{i_r,j_r} with minimum $j_r \geq 3$. Let us call $j_s = \min_{r\in\{1,\cdots,m\}} j_r$. Let us compare $a(d1, l_s)$ and $a(d1, l_o)$ where $j_o > j_s$. Using (5.82), we obtain :

$$
\begin{aligned}
W_{a(d1,l_o)}(3,t) - W_{a(d1,l_s)}(3,t) &= \frac{1}{m}\left(\frac{j_s-2}{j_s} + \frac{1}{p}\mathbb{H}_{p2}\right) + \frac{1}{m}\left(\frac{j_o-3}{j_o} + \frac{1}{p}\mathbb{H}_{p2}\right) \\
&- \frac{1}{m}\left(\frac{j_o-2}{j_o} + \frac{1}{p}\mathbb{H}_{p2}\right) - \frac{1}{m}\left(\frac{j_s-3}{j_s} + \frac{1}{p}\mathbb{H}_{p2}\right) \\
&= \frac{1}{j_s} - \frac{1}{j_o} \\
&> 0 \text{ for all } j_s < j_o.
\end{aligned}
\tag{5.86}
$$

Between a $j_r \geq 3$ and $j_o = 2$, difference in value of (5.82) under these two actions is :

$$\begin{aligned}
W_{a(d1,l_2)}(3,t) - W_{a(d1,l_{j_r})}(3,t) &= \frac{1}{m}\left(\frac{j_r-2}{j_r} + \frac{1}{p}\mathbb{H}_{p2}\right) + \frac{1}{mp^2}Tr(2:1,0) \\
&\quad - \left(\frac{1}{mp}\mathbb{H}_{p2} + \frac{1}{m}\frac{j_r-3}{j_r} + \frac{1}{mp}\mathbb{H}_{p2}\right) \\
&= \frac{1}{mj_r} + \frac{1}{mp^2}\,Tr(2:1,0) - \frac{1}{mp}\mathbb{H}_{p2} \qquad (5.87) \\
&= \frac{1}{mj_r} + \frac{1}{mp^2}\left(\frac{5p}{2} + \frac{7}{4} - \frac{5}{2}\mathbb{H}_p - p\mathbb{H}_p\right) \\
&> 0.
\end{aligned}$$

We obtain the inequality from (5.73). This implies that it is better to mark the leaf of subtree t_{i_r,j_r} with $j_r \geq 3$.
If all the $j_r < 3$, then let us check whether it is better to mark leaf with one sibling or none. Assume there is a $j_r = 1$ and $j_s = 2$ for some $r, s \in [1, m]$.

$$\begin{aligned}
W_{a(d1,l_2)}(3,t) - W_{a(d1,l_1)}(3,t) &= \frac{1}{mp}\mathbb{H}_{p3} + \frac{1}{mp^2}Tr(2:1,0) - \left(\frac{1}{mp}\mathbb{H}_{p4} + \frac{1}{mp}\mathbb{H}_{p2}\right) \\
&= \frac{1}{mp}\left(3\mathbb{H}_p - p - \frac{17}{6}\right) + \frac{1}{mp^2}\,Tr(2:1,0) \\
&= \frac{1}{mp^2}\left(\frac{2p}{3} + \frac{7}{4} - \frac{5}{2}\mathbb{H}_p\right) \\
&< 0.
\end{aligned} \qquad (5.88)$$

This means that it is better to mark the leaf with one sibling.
Hence, completing the statement of Lemma 5.7, we have :

Lemma 5.8. *It is better to mark a leaf of the subtree with* $\min_r\{j_r \geq 3\}$ *if any* $j_r \geq 3$*, else we mark leaf of subtree with* $j_r = 2$ *if it exists, else we mark leaf of the subtree with* $j_r = 1$.

Before moving to Type $2b$, let us solve the trivial case of Type $2c$.

Type 2c :
The only action here is marking two leaves and hence the value function is :

$$W_{a(l_{j_1},l_{j_1})}(3,t) = \left(\mathbb{1}_{\{j_1\geq 4\}}\left(\frac{(j_1-4)^+}{j_1} + \frac{1}{p}\mathbb{H}_{p2}\right)\right) + \frac{1}{p^3}\cdot(\mathbb{1}_{\{j_1=3\}})Tr(3:1,1,0)$$
$$+ \frac{1}{p^2}\cdot(\mathbb{1}_{\{j_1=2\}})Tr(2:1,1) + \frac{1}{p}\cdot(\mathbb{1}_{\{j_1=1\}})\sum_{f_1=1}^{p}\frac{(f_1-4)^+}{f_1}. \quad (5.89)$$

Type 2b :

The action set is : $A = \{a(l_{j_1}, l_{j_2}), a(l_{j_1}, l_{j_1}), a(l_{j_2}, l_{j_2})\}$. We evaluate the value function under these actions. Under action of marking two leaves with different fathers :

$$\begin{aligned}
W_{a(l_{j_1},l_{j_2})}(3,t) &= c(t, a(l_{j_1}, l_{j_2})(t)) + \sum_{t'\in\mathcal{SD}(t)} P(t, a(l_{j_1}, l_{j_2})(t), t')\cdot W^*(2,t') \\
&= 0 + \sum_{r=1}^{2}\frac{1}{2p^{j_r}}\sum_{t'\in\mathcal{D}(t_{i_r,j_r})} W^*(2,t') \\
&= \frac{1}{2p^{j_1}}\sum_{1\leq f_1,\cdots,f_{j_1}\leq p} W^*(2,(\mu,\{t_{1,f_1},\cdots,t_{0,f_{j_1}}\})) \\
&\quad + \frac{1}{2p^{j_2}}\sum_{1\leq f_1,\cdots,f_{j_2}\leq p} W^*(2,(\mu,\{t_{1,f_1},\cdots,t_{0,f_{j_2}}\})),
\end{aligned} \quad (5.90)$$

Splitting individual cases of j_1, j_2 using indicators,

$$\begin{aligned}
W_{a(l_{j_1},l_{j_2})}(3,t) &= \frac{1}{2}\cdot\left(\mathbb{1}_{\{j_1\geq 3\}}\left(\frac{(j_1-3)}{j_1} + \frac{1}{p}\mathbb{H}_{p2}\right)\right) \\
&\quad + \frac{1}{2}\cdot\left(\mathbb{1}_{\{j_2\geq 3\}}\left(\frac{(j_2-3)}{j_2} + \frac{1}{p}\mathbb{H}_{p2}\right)\right) \\
&\quad + \frac{1}{2p^2}\cdot(\mathbb{1}_{\{j_1=2\}} + \mathbb{1}_{\{j_2=2\}})Tr(2:1,0) \\
&\quad + \frac{1}{2p}\cdot(\mathbb{1}_{\{j_1=1\}} + \mathbb{1}_{\{j_2=1\}})\sum_{f_1=1}^{p}\frac{(f_1-4)^+}{f_1}.
\end{aligned} \quad (5.91)$$

Similarly for the action $a(l_c, l_c)$ where $c \in \{1,2\}$ (without loss of generality, assume first two leaves are marked), we have:

$$W_{a(l_c,l_c)}(3,t) = \frac{1}{2}\cdot\sum_{r\in\{1,2\}/\{c\}}\left(\mathbb{1}_{\{j_r\geq 2\}}\left(\frac{(j_r-2)^+}{j_r} + \frac{1}{p}\mathbb{H}_{p2}\right) + \mathbb{1}_{\{j_r=1\}}\left(\frac{1}{p}\mathbb{H}_{p3}\right)\right)$$

$$+\frac{1}{2}\cdot\left(\mathbb{1}_{\{j_c\geq 4\}}\left(\frac{j_c-4}{j_c}+\frac{1}{p}\mathbb{H}_{p2}\right)\right)+\frac{1}{2p^3}\cdot(\mathbb{1}_{\{j_c=3\}})Tr(3:1,1,0)$$

$$+\frac{1}{2p^2}\cdot(\mathbb{1}_{\{j_c=2\}})Tr(2:1,1)+\frac{1}{2p}\cdot(\mathbb{1}_{\{j_c=1\}})\sum_{f_1=1}^{p}\frac{(f_1-4)^+}{f_1}. \tag{5.92}$$

Let us compare the value equations under the three actions in A. If the difference is positive then $a(l_{j_1},l_{j_1})$ is better than $a(l_{j_1},l_{j_2})$, and vice versa. Therefore,

$$W_{a(l_{j_1},l_{j_2})}(3,t)-W_{a(l_{j_1},l_{j_1})}(3,t)$$

$$=\begin{cases}\frac{1}{2}\left(\frac{1}{j_1}-\frac{1}{j_2}\right) & \text{for } j_1\geq 4, j_2\geq 4 \quad (5.93)\\[2ex] \frac{-1}{2j_2}+\frac{1}{2p}(p-2\mathbb{H}_p+1) & \\ -\frac{1}{2p^3}Tr(3:1,1,0) & \text{for } j_1=3, j_2\geq 3 \quad (5.94)\\[2ex] \frac{1}{2p^2}\cdot Tr(2:1,0)-\frac{1}{2j_2} & \\ -\frac{1}{2p^2}\cdot Tr(2:1,1) & \text{for } j_1=2, j_2\geq 3 \quad (5.95)\\[2ex] \frac{1}{2p^2}\cdot 2Tr(2:1,0)-\frac{1}{2p}(p-2\mathbb{H}_p+1) & \\ -\frac{1}{2p^2}\cdot Tr(2:1,1) & \text{for } j_1=2, j_2=2 \quad (5.96)\\[2ex] \frac{-1}{2j_2} & \text{for } j_1=1, j_2\geq 3 \quad (5.97)\\[2ex] \frac{1}{2p^2}\cdot Tr(2:1,0)-\frac{1}{2p}(p-2\mathbb{H}_p+1) & \text{for } j_1=1, j_2=2 \quad (5.98)\\[2ex] \frac{1}{2p}\left((p-4\mathbb{H}_p+\frac{13}{3})-(p-3\mathbb{H}_p+\frac{5}{2})\right) & \text{for } j_1=1, j_2=1. \quad (5.99)\end{cases}$$

In (5.93), the difference is obviously positive since $j_1 \leq j_2$. (5.94) is negative for some values of j_2 and then switches sign as j_2 increases. The threshold for j_2 after which (5.94) turns positive is :

$$j_2 \geq \frac{p^3}{\mathbb{H}_p(\frac{19}{3} + 5p + p^2) - 3p^2 - \frac{5}{2} - 8p} \tag{5.100}$$

Hence $a(l_{j_1}, l_{j_1})$ is better than $a(l_{j_1}, l_{j_2})$ as j_2 crosses the threshold which can be verified by referring to the results of a numerical simulation in Figure 5.2 (Page 70). The case for $j_1 = 2, j_2 \geq 3$ in (5.95) relies on j_2 and hence gives rise to an interesting phenomenon when simulated numerically.
Given a particular p, (5.95) is negative for some j_2 and after j_2 crosses a "threshold" (5.95) becomes positive. This means the optimal action switches from $a(l_{j_1}, l_{j_2})$ to $a(l_{j_1}, l_{j_1})$ after j_2 crosses a threshold. The threshold is precisely,

$$j_2 \geq \frac{12p^2}{12p\mathbb{H}_p - 34\mathbb{H}_p - 39 + 42\mathbb{H}_p} \tag{5.101}$$

Of course we also impose the condition that $j_2 \leq p$. We can verify this by referring to the numerical simulation in Figure 5.3 (Page 70).

Let us tackle the case for $j_1 = 2, j_2 = 2$ in (5.96).

$$\frac{1}{2p^2} \cdot 2Tr(2:1,0) - \frac{1}{2p}\mathbb{H}_{p2} - \frac{1}{2p^2} \cdot Tr(2:1,1) = \frac{1}{2p^2}\left(\mathbb{H}_p - \frac{p}{3} - \frac{3}{2}\right) < 0. \tag{5.102}$$

(5.97) is trivially negative. Let us simplify (5.98) :

$$\frac{1}{2p^2} \cdot Tr(2:1,0) - \frac{1}{2p}\mathbb{H}_{p2} = \frac{1}{2p^2}\left(\frac{7}{4} - p\mathbb{H}_p + \frac{3p}{2} - \frac{5\mathbb{H}_p}{2}\right) < 0. \tag{5.103}$$

(5.99) is also trivially negative.

Similarly we compare the other pair actions $a(l_{j_2}, l_{j_2}), a(l_{j_1}, l_{j_2})$. If the difference is positive then $a(l_{j_1}, l_{j_2})$ is better than $a(l_{j_2}, l_{j_2})$ and vice versa.

$$W_{a(l_{j_2},l_{j_2})}(3,t) - W_{a(l_{j_1},l_{j_2})}(3,t)$$

$$= \begin{cases} \frac{1}{2}\left(\frac{1}{j_1} - \frac{1}{j_2}\right) & \text{for } j_1 \geq 3, j_2 \geq 4 \quad (5.104) \\ \frac{1}{6} - \frac{1}{2p}(p - 2\mathbb{H}_p + 1) + \frac{1}{2p^3}Tr(3:1,1,0) & \text{for } j_1 = 3, j_2 = 3 \quad (5.105) \\ \frac{1}{2p}(p - 2\mathbb{H}_p + 1) - \frac{1}{2p^2}\cdot Tr(2:1,0) - \frac{1}{2j_2} & \text{for } j_1 = 2, j_2 \geq 4 \quad (5.106) \\ \frac{1}{2p^3}Tr(3:1,1,0) - \frac{1}{2p^2}\cdot 2Tr(2:1,0) & \text{for } j_1 = 2, j_2 = 3 \quad (5.107) \\ \frac{1}{2p}(p - 2\mathbb{H}_p + 1) + \frac{1}{2p^2}\cdot Tr(2:1,1) & \\ -\frac{1}{2p^2}\cdot 2Tr(2:1,0) & \text{for } j_1 = 2, j_2 = 2 \quad (5.108) \\ -\frac{1}{2j_2} + \frac{1}{2p}\left((p - 3\mathbb{H}_p + \frac{5}{2}) - (p - 4\mathbb{H}_p + \frac{13}{3})\right) & \text{for } j_1 = 1, j_2 \geq 4 \quad (5.109) \\ \frac{1}{2p^3}Tr(3:1,1,0) + \frac{1}{2p}\left(p - 3\mathbb{H}_p + \frac{5}{2}\right) & \\ -\frac{1}{2p}\left((p - 4\mathbb{H}_p + \frac{13}{3}) + (p - 2\mathbb{H}_p + 1)\right) & \text{for } j_1 = 1, j_2 = 3 \quad (5.110) \\ \frac{1}{2p}\left((p - 3\mathbb{H}_p + \frac{5}{2}) - (p - 4\mathbb{H}_p + \frac{13}{3})\right) & \\ +\frac{1}{2p^2}\left(Tr(2:1,1) - Tr(2:1,0)\right) & \text{for } j_1 = 1, j_2 = 2 \quad (5.111) \\ \frac{1}{2p}\left((p - 3\mathbb{H}_p + \frac{5}{2}) - (p - 4\mathbb{H}_p + \frac{13}{3})\right) & \text{for } j_1 = 1, j_2 = 1. \quad (5.112) \end{cases}$$

(5.104) is obviously non-negative since $j_2 \geq j_1$. (5.105) is positive since it is the negative of (5.94) for $j_2 = 3$. (5.106) is negative for some j_2 values and changes sign after j_2 crosses a threshold. Precisely, (5.106) is positive when :

$$j_2 \geq \frac{4p^2}{4p\mathbb{H}_p - 10p - 7 + 10\mathbb{H}_p} \tag{5.113}$$

Of course we also impose the condition that $j_2 \leq p$. This too has been verified numerically as evident from Figure 5.3 (Page 70). (5.107) simplifies to

$$\frac{1}{2p^3}\left(\frac{5}{2} + \frac{9p}{2} + 3p^2\mathbb{H}_p - 3p^2 - p^3 - \frac{19\mathbb{H}_p}{3}\right) < 0. \tag{5.114}$$

(5.108) is the negative of (5.96) and hence is positive.

(5.109) depends on j_2 and is positive when :

$$j_2 \geq \frac{p}{\mathbb{H}_p - \frac{11}{6}} \tag{5.115}$$

This means $a(l_{j_1}, l_{j_2})$ is optimal when j_2 crosses the threshold. Figure 5.4 (Page 71) depicts this switch between optimal policies.
(5.110) evaluates to :

$$\frac{1}{2p^3}\left(\frac{5}{2} + 8p + \frac{7p^2}{6} - \frac{1}{3}(19 + 15p)\mathbb{H}_p\right) > 0 \text{ for } p \geq 5. \tag{5.116}$$

(5.111) simplifies to :

$$\begin{aligned}&\frac{1}{2p}\left(\mathbb{H}_{p3} - \mathbb{H}_{p4}\right) + \frac{1}{2p^2}\left(Tr(2:1,1) - Tr(2:1,0)\right)\\ &= \frac{1}{2p^2}\left(p + \frac{13}{4} - \frac{7\mathbb{H}_p}{2}\right) > 0.\end{aligned} \tag{5.117}$$

(5.112) is trivially positive. Figure 5.2 (Page 70) depicts the optimal policies for type $2b$ trees with $j_1 = 3$ for different values of p, j_2. The Type $2b$ tree with $j_1 = 2, j_2 > 3$ is a special case where all three actions are optimal for some configuration of p, j_2. Figure 5.3 (Page 70) depicts the dynamics of the optimal actions as p varies in the interval $[4, 50]$ on the x-axis, and j_2 varies in the interval $[4, p]$ on the y-axis. The coloured bubbles represent different optimal actions according to the legend. To read the plot, pick a point p on the x-axis.

Then go vertically up along the values of j_2 along the y-axis. The colour of the bubble indicates the optimal policy for a particular p and $3 < j_2 \leq p$.

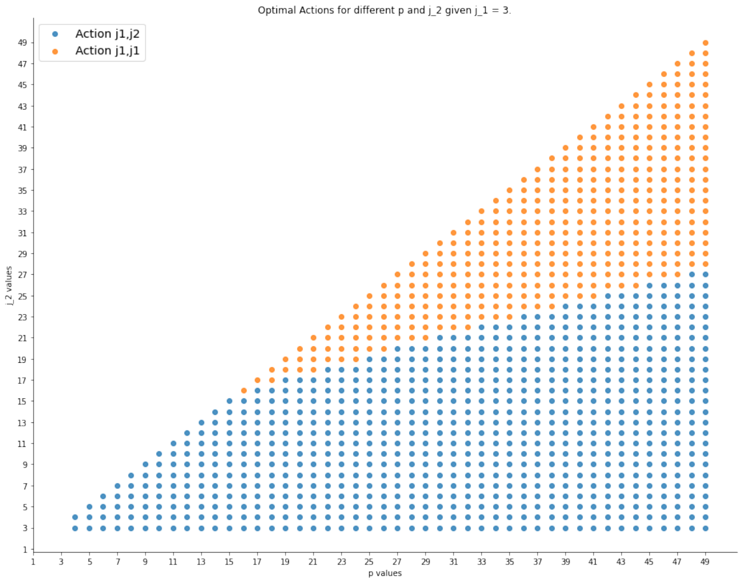

Figure 5.2: Optimal Actions for different values of p and j_2 given $j_1 = 3, j_2 \geq 3$

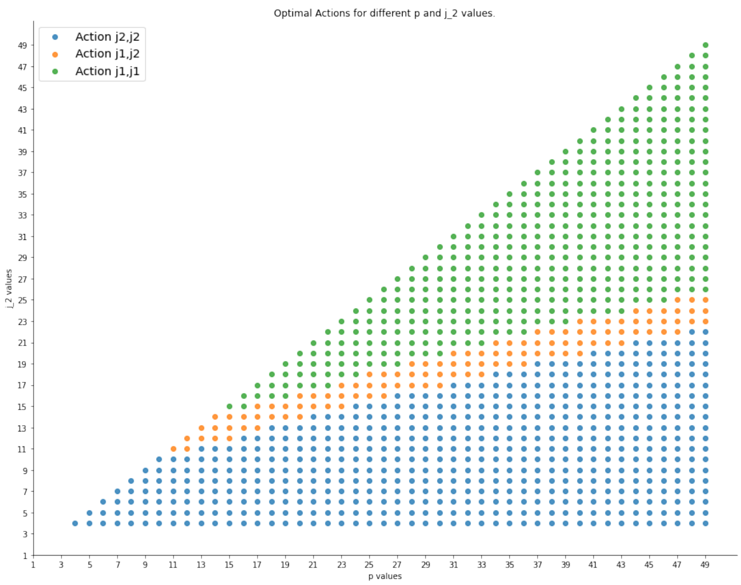

Figure 5.3: Optimal Actions for different values of p and j_2 given $j_1 = 2, j_2 > 3$

Just like for $j_1 = 2, 3$, Figure 5.4 depicts the different optimal policies when $j_1 = 1$ for different values of p, j_2.

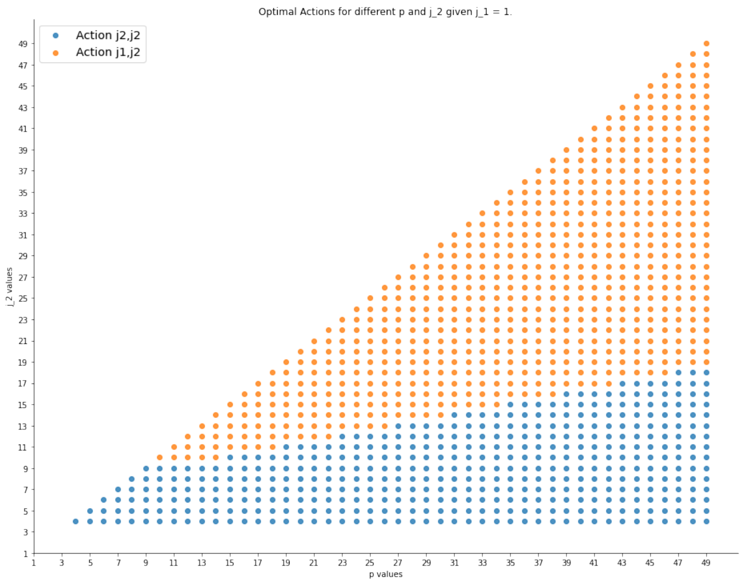

Figure 5.4: Optimal Actions for different values of p and j_2 given $j_1 = 1, j_2 > 3$

The table below summarises the optimal actions for all tree types at $n = 3$.

Table 5.5: Optimal Actions for Trees at $n = 3$

Tree Type	Specifications		Optimal Action
Type 1			$a(d1, d1)$
Type 2a	$j_r \geq 3$ for some r		$a(d1, l_{j_c})$, $j_c = \min_r j_r \geq 3$
	$j_r < 3$ for all r		$a(d1, l_{j_c})$, $j_c = \max_r j_r$
Type 2b	$j_1 > 3$	$j_2 > 3$	$a(l_{j_1}, l_{j_1})$ assuming $j_1 \leq j_2$
	$j_1 = 3$	$j_2 \geq 3$	$a(l_{j_1}, l_{j_1})$ and $a(l_{j_1}, l_{j_2})$ determined by the threshold in (5.100).
	$j_1 = 2$	$j_2 > 3$	$a(l_{j_1}, l_{j_1})$, $a(l_{j_1}, l_{j_2})$, and $a(l_{j_2}, l_{j_2})$ determined by thresholds in (5.101) and (5.113).
		$j_2 = 3$	$a(l_{j_2}, l_{j_2})$
		$j_2 = 2$	$a(l_{j_1}, l_{j_2})$
	$j_1 = 1$	$j_2 \geq 4$	$a(l_{j_2}, l_{j_2})$ and $a(l_{j_1}, l_{j_2})$ determined by the threshold in (5.115)
		$j_2 = 3$	$a(l_{j_2}, l_{j_2})$ for $p < 5$, $a(l_{j_1}, l_{j_2})$ for $p \geq 5$
		$j_2 < 3$	$a(l_{j_1}, l_{j_2})$
Type 2c			$a(l_{j_1}, l_{j_1})$

5.2.4 Horizon $n = 4$

We move to the horizon $n = 4$. The $n = 3$ case had complicated value equations. Surely the $n = 4$ case would contain added complexity. Hence, we numerically analyse the value functions for $p \in [1, 500]$ to infer the optimal actions. All Type 1 trees were investigated.
Table 5.6 summarises the optimal actions for all tree types at $n = 4$.

Table 5.6: Optimal Actions for Trees at $n = 4$

Tree Type	Specifications		Optimal Action
Type 1			$a(d1, d1)$
Type 2a	$j_r \geq 3$ for some r		$a(d1, l_{j_c})$, $j_c = \min_r j_r \geq 3$
	$j_r < 3$ for all r		$a(d1, l_{j_c})$, $j_c = \max_r j_r$
Type 2b	$j_1 > 3$	$j_2 > 3$	$a(l_{j_1}, l_{j_1})$
	$j_1 = 3$	$j_2 \geq 3$	Switch from $a(l_{j_1}, l_{j_2})$ to $a(l_{j_1}, l_{j_1})$ after $p \geq 13$ for some j_2.
	$j_1 = 2$	$j_2 > 3$	Switch from $a(l_{j_1}, l_{j_2})$ to $a(l_{j_1}, l_{j_1})$ after $p \geq 52$ for some j_2.
		$j_2 = 3$	$a(l_{j_1}, l_{j_2})$
		$j_2 = 2$	$a(l_{j_1}, l_{j_2})$
	$j_1 = 1$	$j_2 > 3$	$a(l_{j_1}, l_{j_2})$ mostly, but $a(l_{j_2}, l_{j_2})$ for some values of j_2 dependent on p.
		$j_2 < 3$	$a(l_{j_1}, l_{j_2})$
Type 2c			$a(l_{j_1}, l_{j_1})$

It is interesting to note how the optimal actions have changed (or not) from $n = 3$ to $n = 4$. Table 5.7 summarises the changes for Type $2b$ trees since they are of interest.

Table 5.7: Optimal actions for Type 2b trees at $n = 3, 4$

Sl no.	Type 2b Tree Shape	$n = 3$	$n = 4$
1	$j_1, j_2 > 3$	$a(l_{j_1}, l_{j_1})$	$a(l_{j_1}, l_{j_1})$
2	$j_1 = 3, j_2 \geq 3$	Switch from $a(l_{j_1}, l_{j_2})$ to $a(l_{j_1}, l_{j_1})$ determined by the threshold in (5.100)	Switch from $a(l_{j_1}, l_{j_2})$ to $a(l_{j_1}, l_{j_1})$ after $p \geq 13$ for some j_2.
3	$j_1 = 2, j_2 > 3$	Switch from $a(l_{j_2}, l_{j_2})$ to $a(l_{j_1}, l_{j_2})$ to $a(l_{j_1}, l_{j_1})$, determined by thresholds in (5.101) and (5.113)	Switch from $a(l_{j_1}, l_{j_2})$ to $a(l_{j_1}, l_{j_1})$ after $p \geq 52$ for some j_2.
4	$j_1 = 2, j_2 = 3$	$a(l_{j_2}, l_{j_2})$	$a(l_{j_1}, l_{j_2})$
5	$j_1 = 2, j_2 = 2$	$a(l_{j_1}, l_{j_2})$	$a(l_{j_1}, l_{j_2})$
6	$j_1 = 1, j_2 > 3$	Switch from $a(l_{j_2}, l_{j_2})$ to $a(l_{j_1}, l_{j_2})$ determined by the threshold in (5.115)	$a(l_{j_1}, l_{j_2})$ mostly, but $a(l_{j_2}, l_{j_2})$ for some values of j_2 dependent on p.
7	$j_1 = 1, j_2 = 3$	$a(l_{j_2}, l_{j_2})$ for $p < 5$, $a(l_{j_1}, l_{j_2})$ for $p \geq 5$	$a(l_{j_1}, l_{j_2})$
8	$j_1 = 1, j_2 < 3$	$a(l_{j_1}, l_{j_2})$	$a(l_{j_1}, l_{j_2})$

It is evident from Table 5.7 that the optimal actions at $n = 4$ are relatively simpler than at $n = 3$. This give us reason to believe that as n increases, the optimal actions may not be as varied for different sizes of subtrees (i.e, different j_1, j_2 values). Moreover, the thresholds may disappear after some large value of n. We note that this specific greedy policy at $n = 4$, which we shall refer to as the "Leg-1 Policy" is a candidate policy for the infinite horizon criterion. The complexity of numerical analysis for $n = 5$ and greater is out of scope of this study, and away from the aim of this thesis. We shall truncate the finite horizon criterion here and move to the infinite horizon average costs criterion.

Chapter 6

Infinite Horizon Average Costs Prefetching

The infinite horizon average costs criterion has an interesting feature. We only have to look among stationary policies in our search for an optimal policy. Just like in the finite horizon case, we shall first establish some results for depth 1 trees and then move to depth 2 trees.
Before we get to these specific depths, it is important to first eliminate "wasteful" actions, such as those which mark the root node, or already marked nodes, or not using the complete budget. To do so, we shall prove that the action of marking a root node is always dominated by some action that does not. Note that marking the root is equivalent to marking an already marked node, or not using one mark at all.

6.1 Eliminating wasteful actions

For a general p, d, let $\tilde{\pi}$ be the policy of marking the root and $k-1$ sons. Recall the representation of a marked tree of depth d as $t = (\mu, s)$ where s is a list of trees of depth $d-1$. Let us eliminate the action of marking the root for each of the two types of trees : trees with number of unmarked sons greater than budget and those with number of unmarked sons smaller than the budget. Call the set of trees of the former type as T_{kbig}, and those of the latter type as T_{ksmall}. All notation introduced for both cases has been specified in the List of Notation 2 (23-27). Before moving to the proof, we state and prove a useful lemma.

Recall the notation : For any tree t and action a, $a(t)$ denotes the tree t after marking according to the action a. Similarly, for a stationary policy π, $\pi(t)$ is the tree t after marking according to the action specified by π. Moreover for a tree $t = (\mu, s)$, let $a(s)$ be the "list" of sub-trees of t after marking according to the action a.

Lemma 6.1. *For any tree t, consider any two actions a' and a'' and any $t' \in \mathcal{SD}(a'(t))$, and $t'' \in \mathcal{SD}(a''(t))$ such that t' and t'' have the same shape (i.e, t' and t'' differ only in markings of the nodes). Then $P(t, a', t') = P(t, a'', t'')$.*

Proof. Let us revisit the transition probability structure,

$$P(t, a, t') = \begin{cases} \dfrac{1}{|s||\mathcal{D}(t_b)|} & \text{if } t' \in \mathcal{D}(t_b) \text{ where } t_b \in a(s) \\ \\ 0 & \text{if } t' \notin \mathcal{SD}(a(t)). \end{cases} \tag{6.1}$$

In (6.1), the term $|s|$ in the denominator will be the same for both transitions to t' and t'' since $|s|$ is dependent only on t. We have to show that $|\mathcal{D}(t_b)|$ is the same for both transitions. We only have to show that the size of the sets are equal.
We note the following: Let t'_b and t''_b be the subtrees of $a'(t)$ and $a''(t)$ whose discoveries lead to t' and t'' respectively.

- t'_b and t''_b are sub-trees of t after the actions a' and a'' respectively.
- The shapes of t'_b and t''_b are the same and they differ only in the marking of nodes.
- Since the number of leaves of t'_b and t''_b are equal, the number of discoveries of both the trees are the same.

Hence, when $t'_b \in a'(s)$ and $t''_b \in a''(s)$, we have that $|\mathcal{D}(t'_b)| = |\mathcal{D}(t''_b)|$. Thus $P(t, a', t') = P(t, a'', t'')$ when the shapes of t', t'' are the same. $\square$

Theorem 6.2. *For the prefetching infinite horizon MDP with the state space $\mathcal{T}_{p,d}$, and given any budget k, the action of marking the root node of any tree is always dominated by some other action.*

For simplicity, we shall prove a less general statement first which can be extended to Theorem 6.2.

Theorem 6.3. *For the prefetching infinite horizon MDP with the state space $\mathcal{T}_{p,d}$, and given a budget k; the policy of marking the root first and then $k-1$ unmarked sons for trees in T_{kbig}, and marking the root first with all the unmarked sons and some other nodes for trees in T_{ksmall}, is dominated by another policy which does not mark the root.*

Proof. For each case of trees, we shall define two policies and calculate the difference in values of trees under the two policies. If the differences are of same sign for every n, then dividing the difference by n and taking $n \to \infty$ would prove that one policy dominates the other.

Case 1 : T_{kbig}

Consider two policies π and $\tilde{\pi}$. Let π mark k unmarked sons and $\tilde{\pi}$ mark the root and $k-1$ unmarked sons of trees in T_{kbig}. For trees in T_{ksmall}, let π and $\tilde{\pi}$ do the exact same marking. We are not interested in what that marking is exactly as long as it results in the same costs. Hence,

$$W_\pi(1,t) - W_{\tilde{\pi}}(1,t) = \begin{cases} \dfrac{-1}{|s|} & \text{for all } t \in T_{kbig} \\ 0 & \text{for } t \in T_{ksmall}. \end{cases}$$

As a consequence, we have:

$$W_\pi(1,t) - W_{\tilde{\pi}}(1,t) \leq 0 \qquad \text{for all } t. \tag{6.2}$$

Next, for $n = 2$, for all $t \in T_{kbig}$,

$$\begin{aligned} W_\pi(2,t) - W_{\tilde{\pi}}(2,t) = \frac{-1}{|s|} + & \sum_{t' \in \mathcal{SD}(\pi(t))} P(t,\pi(t),t') W_\pi(1,t') \\ - & \sum_{t' \in \mathcal{SD}(\tilde{\pi}(t))} P(t,\tilde{\pi}(t),t') W_{\tilde{\pi}}(1,t'). \end{aligned}$$

There is one subtree (call it t_{rm}) whose discovery would have a marked root under π, but an unmarked root under $\tilde{\pi}$. Thus, for the set of trees $\{t' : t' \in \mathcal{SD}(\pi(t))/\mathcal{D}(t_{rm})\}$, there is a corresponding set of trees $\{\tilde{t'} : \tilde{t'} \in \mathcal{SD}(\tilde{\pi}(t))/\mathcal{D}(t_{rm})\}$ with the same shape and marking.

The trees $t'_m \in \mathcal{D}(t_{rm})$ and $t'_{un} \in \mathcal{D}(t_{rm})$ have the shape same, and hence using Lemma 6.1, we have that for all $t \in T_{kbig}$:

$$\begin{aligned} W_\pi(2,t) - W_{\tilde{\pi}}(2,t) = \frac{-1}{|s|} + & \sum_{t' \in \mathcal{SD}(\pi(t))/\mathcal{D}(t_{rm})} P(t,\pi(t),t') \left(W_\pi(1,t') - W_{\tilde{\pi}}(1,t')\right) \\ + & \sum_{t' \in \mathcal{D}(t_{rm})} P(t,\pi(t),t') \left(W_\pi(1,t'_m) - W_{\tilde{\pi}}(1,t'_{un})\right). \end{aligned} \tag{6.3}$$

In (6.3), t'_m is the tree in $\mathcal{D}(t_{rm})$ with marked root, whereas t'_{un} is the tree in $\mathcal{D}(t_{rm})$ with unmarked root.
Since the mark of the root is irrelevant in calculating the values of the trees, the second summation can be absorbed into the first summation and the result is also non-positive. (6.3) is the sum of a negative term and two summations over non-positive terms which would lead to negative difference.

Similarly for $t \in T_{ksmall}$,

$$W_\pi(2,t) - W_{\tilde{\pi}}(2,t) = \sum_{t' \in \mathcal{SD}(\pi(t))} P(t,\pi(t),t') \left(W_\pi(1,t') - W_{\tilde{\pi}}(1,t')\right). \tag{6.4}$$

The difference in (6.4) is obviously non-positive from (6.2). The immediate costs are the same since the policies are the same for T_{ksmall} trees. The difference in future costs is non-positive. We have therefore,

$$W_\pi(2,t) - W_{\tilde{\pi}}(2,t) \leq 0 \qquad \forall t \in \mathcal{T}_{d,p}.$$

By induction, we can prove that:

$$W_\pi(n,t) - W_{\tilde{\pi}}(n,t) \leq 0 \qquad \forall n = 0,1,\ldots,N \quad \forall t \in \mathcal{T}_{d,p}.$$

This is indeed true for $n \leq 2$ as proved before, and for general n,

$$W_\pi(n,t) - W_{\tilde{\pi}}(n,t)$$

$$= \begin{cases} \dfrac{-1}{|s|} + \displaystyle\sum_{t' \in \mathcal{SD}(\pi(t))/\mathcal{D}(t_{rm})} P(t,\pi(t),t')\left(W_\pi(n-1,t') - W_{\tilde{\pi}}(n-1,t')\right) & \\ + \displaystyle\sum_{t' \in \mathcal{D}(t_{rm})} P(t,\pi(t),t')\left(W_\pi(n-1,t'_m) - W_{\tilde{\pi}}(n-1,t'_{un})\right) & \text{for all } t \in T_{kbig} \\ & \\ \displaystyle\sum_{t' \in \mathcal{SD}(\pi(t))} \left(W_\pi(n-1,t') - W_{\tilde{\pi}}(n-1,t')\right) & \text{for all } t \in T_{ksmall}. \end{cases} \tag{6.5}$$

Both the expressions in (6.5) will be non-positive because of the non-positive difference in values at $n-1$. Hence, policy π dominates the policy $\tilde{\pi}$.

Case 2 : T_{ksmall}
For trees in T_{ksmall}, there are more subtleties.

- Marking a root would still lead to sufficient budget for marking all unmarked sons plus some more nodes.
- Let $\bar{\pi}$ mark the root, the unmarked sons, and some other nodes in the tree. Let us assume that π does the same markings as $\bar{\pi}$ except that π does not mark the root and instead marks an additional node (call it l) at some depth $\delta(l,t) \in [2,d]$ of a tree t.
- Assume that these two policies do the exact same markings for trees in T_{kbig}.

Trivially we would have that $W_\pi(1,t) - W_{\bar{\pi}}(1,t) = 0$ for all t. There is a non trivial difference at $n = 2$ depending on where the node l is. Let the number of unmarked siblings of l be denoted by $sib(l)$. Of course if $sib(l) \geq k$, discovery of the subtree where l is a son would be a tree belonging to T_{kbig}. Call the $d-1$ subtree where l

is a node as t_l. For the specific case that $t \in T_{ksmall}, \delta(l,t) = 2$, $sib(l) \geq k$,

$$W_\pi(2,t) - W_{\bar{\pi}}(2,t) = \sum_{t' \in \mathcal{SD}(t)/\mathcal{D}(t_l)} P(t,\pi(t),t') \left(W_\pi(1,t') - W_{\bar{\pi}}(1,t')\right)$$
$$+ \sum_{t' \in \mathcal{D}(t_l)} P(t,\pi(t),t') \left(W_\pi(1,t'_{lm}) - W_{\bar{\pi}}(1,t'_{lun})\right) \quad (6.6)$$
$$= \sum_{t' \in \mathcal{D}(t_l)} P(t,\pi(t),t') \frac{-1}{tsib(l)+1}. \quad (6.7)$$

In (6.6), the trees t'_{lm} are the trees in $\mathcal{D}(t_l)$ where node l is marked. Similarly the trees t'_{lun} are the trees in $\mathcal{D}(t_l)$ where node l is unmarked. To take the probability term in common, we use Lemma 6.1. For all other cases of $\delta(l,t)$ and $sib(l)$, and $t \in T_{kbig}$, the difference is zero. We therefore have,

$$W_\pi(2,t) - W_{\bar{\pi}}(2,t) \leq 0 \qquad \forall t \in \mathcal{T}_{d,p}.$$

By induction, we can prove that:

$$W_\pi(n,t) - W_{\bar{\pi}}(n,t) \leq 0 \qquad \forall n = 0,1,\ldots,N \quad \forall t \in \mathcal{T}_{d,p}. \quad (6.8)$$

This is indeed true for $n \leq 2$ as proved above. For $n \geq 3$, the cases where $\delta(l,t) = 2,3,\cdots,\min(n,d)$ are of interest. For $t \in T_{ksmall}$,

$$W_\pi(n,t) - W_{\bar{\pi}}(n,t)$$
$$= \sum_{t' \in \mathcal{SD}(t)/\mathcal{D}(t_l)} P(t,\pi(t),t') \left(W_\pi(n-1,t') - W_{\bar{\pi}}(n-1,t')\right)$$
$$+ \sum_{t' \in \mathcal{D}(t_l)} P(t,\pi(t),t'_{lm}) W_\pi(n-1,t'_{lm}) - \sum_{t' \in \mathcal{D}(t_l)} P(t,\bar{\pi}(t),t'_{lun}) W_{\bar{\pi}}(n-1,t'_{lun}).$$
$$(6.9)$$

In (6.9) the first summation of differences is non-positive according to the induction statement. There is a correspondence between t'_{lm} and t'_{lun} and from Lemma 6.1 the transition probabilities to both these trees are the same.

$$W_\pi(n,t) - W_{\bar\pi}(n,t) = \text{non-positive term}$$

$$+\begin{cases} \sum_{t' \in \mathcal{D}(t_l)} P(t,\pi(t),t') \dfrac{-1}{tsib(l)+1} & \text{if } \delta(l,t)=2,\ sib(l) \ge k \\ \\ \sum_{t' \in \mathcal{D}(t_l)} P(t,\pi(t),t')(W_\pi(n-1,t') & \\ -W_{\bar\pi}(n-1,t')) & \text{if } \delta(l,t)=2,\ sib(l) < k \\ \\ \sum_{t' \in \mathcal{D}(t_l)} P(t,\pi(t),t')(W_\pi(n-1,t'_{lm}) & \\ -W_{\bar\pi}(n-1,t'_{lun})) & \text{if } \delta(l,t) \ge 3. \end{cases} \quad (6.10)$$

The first two cases in (6.10) are non-positive from Lemma 6.1 and the induction statement. For $\delta(l,t) \ge 3$, there are more intricacies to be handled since the benefit of the marked node l is not immediate or even in the first step ahead.

When $\delta(l,t) = 3$, at the next time step, the node l will be at depth 2. To find the sign of $W_\pi(n-1,t'_{lm}) - W_{\bar\pi}(n-1,t'_{lun})$, we can compare each of these two values with $W_\pi(n-1,t'_{lun})$. We can use that for $\delta(l,t'_{lm}) = \delta(l,t'_{lun}) = 2$,

$$W_\pi(n-1,t'_{lm}) - W_\pi(n-1,t'_{lun}) \le 0 \quad (6.11)$$

$$W_\pi(n-1,t'_{lun}) - W_{\bar\pi}(n-1,t'_{lun}) \le 0 \text{ for all } n. \quad (6.12)$$

The second inequality (6.12) follows directly from the induction statement. For the first inequality (6.11),

$$W_\pi(n-1,t_{lm}) - W_\pi(n-1,t_{lun})$$
$$= \sum_{t' \in \mathcal{SD}(t_{lm})} P(t,\pi(t),t')W_\pi(n-2,t') - \sum_{t' \in \mathcal{SD}(t_{lun})} P(t,\pi(t),t')W_\pi(n-2,t') \quad (6.13)$$

$$
= \begin{cases} \sum_{t' \in \mathcal{SD}(t_{lm})} P(t,\pi(t),t') \dfrac{-1}{tsib(l)+1} & \text{if } sib(l) \geq k \text{ and } \delta(t',l) = 1 \\ \\ \sum_{t' \in \mathcal{D}(t_l)} P(t,\pi(t),t')(W_\pi(n-2,t'_{lm}) \\ -W_{\bar{\pi}}(n-2,t'_{lun})) & \text{if } sib(l) < k \text{ and } \delta(t',l) = 1. \end{cases} \tag{6.14}
$$

We have used Lemma 6.1 and the fact that every $t' \in \mathcal{SD}(t_{lm})$ has a corresponding tree $t' \in \mathcal{SD}(t_{lun})$ and the difference in their values would be $-1/(tsib(l)+1)$. For the case where $sib(l) < k$, the difference is non-positive which follows form (6.10) for $n-2$.

If $\delta(l,t)$ is greater than 3, we would need to show that the equations (6.11) and (6.12) hold for any $2 < \delta(l,t') \leq d-1$ where $t' \in \mathcal{D}(t_l)$. This is seen methodically as :

- For $\delta(l,t) = g$ where $g \in [4, d-1]$ for t as in (6.10), we would have to show equations (6.11) and (6.12) hold for $\delta(l,t'_{lm}) = \delta(l,t'_{lun}) = g-1$.
- (6.12) is true from the main induction statement.
- For (6.11), we would proceed by a "depth" induction ; The induction statement would be :

  $$
  W_\pi(n-1,t_{lm}) - W_\pi(n-1,t_{lun}) \leq 0 \text{ for } \delta(l,t_{lm}) = \delta(l,t_{lun}) = g-1 \tag{6.15}
  $$

  where $g \in [4, d-1]$. For the case where $\delta(l,t_{lm}) = \delta(l,t_{lun}) = g$, we would have that

  $$
  \begin{aligned} &W_\pi(n-1,t_{lm}) - W_\pi(n-1,t_{lun}) \\ &= \sum_{t' \in \mathcal{SD}(t_{lm})} P(t,\pi(t),t')W_\pi(n-2,t') - \sum_{t' \in \mathcal{SD}(t_{lun})} P(t,\pi(t),t')W_\pi(n-2,t') \end{aligned} \tag{6.16}
  $$

  where $\delta(l,t') = g-1$.
- The above equation is non-positive according to the induction statement (6.15).

Thus, (6.10) is non positive for all $\delta(l, t)$.
For trees $t \in T_{kbig}$,

$$W_{\pi}(n,t) - W_{\bar{\pi}}(n,t) = \sum_{t' \in \mathcal{SD}(t)} (W_{\pi}(n-1,t') - W_{\bar{\pi}}(n-1,t')) \leq 0. \tag{6.17}$$

This follows from the induction statement for (6.8). Thus policy π is better than the policy $\tilde{\pi}$ in Case 1 for trees in T_{kbig}, and π as defined in Case 2 is better than $\bar{\pi}$ for trees in T_{ksmall}. Combining the two polices for each case would result in a policy that dominates the policy of marking the root of trees. $\square$

6.2 Depth 1 trees

We first define the possible policies for depth 1 trees with a budget k. Instead of directly evaluating the limit $\phi_{\pi}(t)$ as defined in (2.7) for all possible policies π, we shall iteratively compare the values of policies for a general n, and then take the limit of this difference. Later, we compute the stationary distribution of the Markov chain induced by a stationary policy and multiply this stationary distribution vector with the expected cost of each state to obtain the average cost of that policy.

6.2.1 Limit Approach

Label the Usable states of depth 1 trees in U with integers $[1, |U|]$. Since this is the infinite horizon criterion, we look for only stationary policies. There would be mainly two kinds of actions for each tree- Marking root and $k-1$ unmarked sons, or just marking k marked sons. The policies we search over can involve any combination of these two actions for the trees. Let us define the policy:

$$\tilde{\pi} = (\tilde{d}_1, \cdots, \tilde{d}_{|U|}) \tag{6.18}$$

where $\tilde{d}_i$ is the action of marking the root and $k-1$ unmarked sons of the tree labelled i. Let us iteratively define stationary policies $\tilde{\pi}_1, \cdots, \tilde{\pi}_{|U|-1}$ as

$$\tilde{\pi}_1 = (\tilde{d}_1, \cdots, d_{i_1}, \cdots, \tilde{d}_{|U|}) \text{ where } i_1 \in [1, |U|]. \tag{6.19}$$

d_{i_1} is the action of marking k unmarked sons of the tree labelled i_1.

$$\tilde{\pi}_2 = (\tilde{d}_1, \cdots, d_{i_1}, d_{i_2}, \cdots, \tilde{d}_{|U|}) \text{ where } i_1, i_2 \in [1, |U|], i_1 \neq i_2. \tag{6.20}$$

Similarly,

$$\tilde{\pi}_{|U|-1} = (d_{i_1}, \cdots, d_{i_{|U|-1}}, \tilde{d}_{|U|}) \text{ where } i_1, \cdots, i_{|U|-1} \in [1, |U|], i_1 \neq \cdots \neq i_{|U|-1} \tag{6.21}$$

and finally we will have the policy,

$$\tilde{\pi}_{|U|} = \pi = (d_1, \cdots, d_{|U|}) \tag{6.22}$$

Since $i_1, \cdots, i_{|U|}$ is arbitrary, this sequence of policies covers the whole space of policies that we wish to search over. We aim to find the difference $W_{\tilde{\pi}_1}(n, t_{\mu,j}) - W_{\tilde{\pi}}(n, t_{\mu,j})$ and then check if this difference in the limit n going to infinity is negative. To check this, we shall begin with $n = 1, 2, 3$ and then generalise. The policy $\tilde{\pi}_1$ marks the tree t_{μ,i_1} according to the rule d_{i_1}. Thus, for $i_1 \geq k$,

$$W_{\tilde{\pi}_1}(1, t_{\mu,j}) - W_{\tilde{\pi}}(1, t_{\mu,j}) = \begin{cases} 0 & \text{for all } j \neq i_1 \\ \dfrac{-1}{i_1} & \text{for } j = i_1 \end{cases} \tag{6.23}$$

Moving to $n = 2$,

$$W_{\tilde{\pi}_1}(2, t_{\mu,j}) - W_{\tilde{\pi}}(2, t_{\mu,j}) = \begin{cases} \dfrac{-1}{p \cdot i_1} & \text{for all } j \neq i_1 \\ \dfrac{-1}{i_1} - \dfrac{1}{p \cdot i_1} & \text{for } j = i_1 \end{cases} \tag{6.24}$$

When we check for $n = 3$ we identify a pattern emerging,

$$W_{\tilde{\pi}_1}(3, t_{\mu,j}) - W_{\tilde{\pi}}(3, t_{\mu,j}) = \begin{cases} \dfrac{-2}{p \cdot i_1} & \text{for all } j \neq i_1 \\ \dfrac{-1}{i_1} - \dfrac{2}{p \cdot i_1} & \text{for } j = i_1 \end{cases} \tag{6.25}$$

Hence by induction,

$$W_{\tilde{\pi}_1}(n, t_{\mu,j}) - W_{\tilde{\pi}}(n, t_{\mu,j}) = \begin{cases} \dfrac{-(n-1)}{p \cdot i_1} & \text{for all } j \neq i_1 \\ \dfrac{-1}{i_1} - \dfrac{(n-1)}{p \cdot i_1} & \text{for } j = i_1 \end{cases} \tag{6.26}$$

Therefore,

$$\lim_{n\to\infty} \frac{W_{\tilde{\pi}_1}(n, t_{\mu,j}) - W_{\tilde{\pi}}(n, t_{\mu,j})}{n} = \frac{-1}{p \cdot i_1} \text{ for all } t_{\mu,j} \in U. \tag{6.27}$$

Thus, the policy $\tilde{\pi}_1$ is better of than $\tilde{\pi}$ in the average costs criterion.

Comparing the next pair of policies $\tilde{\pi}_1$ and $\tilde{\pi}_2$, we would obtain the same differences but with i_2 instead of i_1. We assume that the same i_1 undergoes the rule of marking k unmarked sons in both the policies.

$$W_{\tilde{\pi}_2}(1, t_{\mu,j}) - W_{\tilde{\pi}_1}(1, t_{\mu,j}) = \begin{cases} 0 & \text{for all } j \neq i_2 \\ \dfrac{-1}{i_2} & \text{for } j = i_2 \end{cases} \tag{6.28}$$

Moving to $n = 2$,

$$W_{\tilde{\pi}_2}(2, t_{\mu,j}) - W_{\tilde{\pi}_1}(2, t_{\mu,j}) = \begin{cases} \dfrac{-1}{p \cdot i_2} & \text{for all } j \neq i_2 \\ \dfrac{-1}{i_2} - \dfrac{1}{p \cdot i_2} & \text{for } j = i_2 \end{cases} \tag{6.29}$$

And by induction we obtain,

$$W_{\tilde{\pi}_2}(n, t_{\mu,j}) - W_{\tilde{\pi}_1}(n, t_{\mu,j}) = \begin{cases} \dfrac{-(n-1)}{p \cdot i_2} & \text{for all } j \neq i_2 \\ \dfrac{-1}{i_2} - \dfrac{(n-1)}{p \cdot i_2} & \text{for } j = i_2 \end{cases} \tag{6.30}$$

Therefore,

$$\lim_{n\to\infty} \frac{W_{\tilde{\pi}_2}(1, t_{\mu,j}) - W_{\tilde{\pi}_1}(1, t_{\mu,j})}{n} = \frac{-1}{p \cdot i_2} \tag{6.31}$$

Thus, we have that for any pair of policies $\tilde{\pi}_q$ with $\tilde{\pi}_{q+1}$ where $q \in [1, |U|]$:

$$W_{\tilde{\pi}_{q+1}}(n, t_{\mu,j}) - W_{\tilde{\pi}_q}(n, t_{\mu,j}) = \begin{cases} \dfrac{-(n-1)}{p \cdot i_{q+1}} & \text{for all } j \neq i_{q+1} \\ \dfrac{-1}{i_{q+1}} - \dfrac{(n-1)}{p \cdot i_{q+1}} & \text{for } j = i_{q+1}. \end{cases} \tag{6.32}$$

Which would lead to the limit of their difference being :

$$\lim_{n\to\infty} \frac{W_{\tilde{\pi}_{q+1}}(n, t_{\mu,j}) - W_{\tilde{\pi}_q}(n, t_{\mu,j})}{n} = \frac{-1}{p \cdot i_{q+1}}. \tag{6.33}$$

Hence, the policy $\tilde{\pi}_{q+1}$ is better of than $\tilde{\pi}_q$ for all $q \in [1, |U|]$ (of course here $|U| = p$). This implies that the policy $\tilde{\pi}_{|U|} = \pi$ is the optimal policy since the policies $\tilde{\pi}_q$ covers all possible markings of all trees. Let us recall the average cost of the policy π from Proposition 4.4.

$$\phi^*(t_{i,j}) = \phi_\pi(t_{i,j}) = \lim_{N\to\infty} \frac{W_\pi(N, t_{i,j})}{N} = \frac{\mathbb{H}_{pk}}{p} \tag{6.34}$$

for all $t_{i,j}$.

Formally, let us refer to this policy of marking sons only as the "Greedy Depth 1" policy.

6.2.2 Stationary Probability Approach

Under policy π, the Usable state space U would consist of trees $t_{0,1}, \cdots, t_{0,p}, \cdots, t_{1,p}$ where $t_{0,j}$ are unmarked trees with j sons, and $t_{1,j}$ are marked trees (marked at the root) with j sons. However, we can aggregate the states to $t_{\mu,1}, \cdots, t_{\mu,p}$ since the mark of the root is irrelevant under the policy we choose to evaluate. The transition structure of this aggregated state space would be :

$$P = \begin{pmatrix} 1/p & 1/p & \cdots & 1/p \\ 1/p & 1/p & \cdots & 1/p \\ \vdots & \vdots & \cdots & \vdots \\ 1/p & 1/p & \cdots & 1/p \\ 1/p & 1/p & \cdots & 1/p \end{pmatrix} \tag{6.35}$$

which would lead to the stationary distribution being,

$$\rho = \begin{pmatrix} 1/p & 1/p & \cdots & 1/p \end{pmatrix}. \tag{6.36}$$

The expected costs for each tree is :

$$E[c(t_{\mu,j}, \pi(t_{\mu,j}))] = \frac{(j-k)^+}{j} \tag{6.37}$$

Therefore, the average cost of policy π is :

$$\sum_{j=1}^{p} \frac{1}{p} E[c(t_{\mu,j}, \pi(t_{\mu,j}))] = \sum_{j=1}^{p} \frac{(j-k)^+}{jp} = \frac{\mathbb{H}_{pk}}{p} = \frac{p-k+1-k(\mathbb{H}_p - \mathbb{H}_{k-1})}{p}. \tag{6.38}$$

This matches with the value of $\phi^*(t)$ calculated using the limit approach.

6.3 Depth 2 trees with budget 1

Given a tree $t = (\mu, \{t_{0,j_1}, \cdots, t_{0,j_m}\})$, the transition structure under policy π is:

$$P(t, \pi(t), t') = \begin{cases} \dfrac{1}{mp^{j_r}} & \text{if } t' \in \mathcal{SD}(t) \\ 0 & \text{otherwise} \end{cases} \tag{6.39}$$

where $\mathcal{SD}(t)$ is the set mapping defined in Definition 3.4.2. Let us divide the state space into blocks. In each block, trees would contain a particular number of sons. There would be p blocks. The first block contains p trees, the second contains p^2, and so on upto p^p trees in the p^{th} block. We call each of these blocks as pm where $m \in [1, p]$ and the block pm contains p^m trees.
The total number of trees are $p + p^2 + \cdots + p^p$. The stationary distribution for the case $p = 3$ is :

$$\rho = \begin{pmatrix} 1/9 & 1/9 & 1/9 & 1/27 & 1/27 & \cdots & 1/27 & 1/81 & \cdots 1/81 \end{pmatrix} \tag{6.40}$$

where $1/27$ repeats for 9 trees in the $p2$ block, and $1/81$ repeats for 27 trees in the $p3$ block.
Using this structure, for a general p, we conjecture that the stationary distribution is :

$$\rho = \begin{pmatrix} 1/p^2 & \cdots & 1/p^2 & 1/p^3 \cdots & 1/p^3 & \cdots & 1/p^{p+1} & \cdots & 1/p^{p+1} \end{pmatrix} \tag{6.41}$$

where the stationary probability of a tree in the pm block is $1/p^{m+1}$ with $m \in [1, p]$. Essentially the trees in each block adds up to a probability of $1/p$, and with p such blocks, the probability of all trees adds to 1. We prove this in the following theorem.

Theorem 6.4. *The stationary distribution under the greedy policy π for depth 2 trees with budget 1 is :*

$$\rho = \begin{pmatrix} 1/p^2 & \cdots & 1/p^2 & 1/p^3 \cdots & 1/p^3 & \cdots & 1/p^{p+1} & \cdots & 1/p^{p+1} \end{pmatrix} \tag{6.42}$$

Proof. We need to check that $\rho P = \rho$. To do so, we shall try to identify the columns in P. We only need to identify the column in P for each block. Let us denote $c(pm)$ as the column vector of transition probabilities for a tree in the pm block.

Therefore $P = (c(p1)c(p2)\cdots c(pp))$, where the column $c(p1)$ repeats p^1 times as it is representative of a tree in the $p1$ block. Similarly the column $c(pm)$ repeats p^m times for every $m \in [1, p]$. Each column is the probability of arriving at a particular tree from all other trees.

Let us breakdown the column vector of transition probabilities of a tree $t = (\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\})$ in the pm block.

- Only one of the trees from the $p1$ block would transition to t. This is so because the number of sons of t will match the number of leaves of exactly one tree in the $p1$ block.

- For the $p2$ block, there are more possibilities. Trees in the $p2$ block with any one or both subtrees having m leaves would have some non zero value. If any one of the subtrees have m leaves, then $1/(2p^m)$ would be the probability of transition from a tree in block pm to our tree of interest. If both subtrees have m leaves, then we simply have the value to be $1/p^m$.

- From any block b, we should have the term $\frac{b-r}{bp^m}$ where $r \in [1, b-1]$ repeats for qr trees. qr represents the number of trees with b sons and exactly $b - r$ subtrees with m leaves. In addition to these terms we should also have the $1/p^m$ term as well from a tree where all the sons have m leaves. And of course there would be some number of $0s$ to complete the column written below.

$$c(pm)^T_b = \begin{pmatrix} \frac{1}{p^m} & \frac{b-1}{bp^m}_{\;q1} & \frac{b-2}{bp^m}_{\;q2} & \frac{b-3}{bp^m}_{\;q3} & \cdots & \frac{b-(b-1)}{bp^m}_{\;q(b-1)} & 0_{qb} \end{pmatrix} \tag{6.43}$$

Let us find qr which is the number of trees with b sons and all but r of the b subtrees have m leaves. First pick r nodes among the b sons, and each of these r nodes can have any one of the $p-1$ possible sons of their own. Thus, qr is simply,

$$qr = \binom{b}{r}(p-1)^r \text{ for all } r \in [0, b]. \tag{6.44}$$

For $r = 0$ (6.44) is simply 1, and for $r = b$ it is $(p-1)^b$.
We evaluate the product ρP. This product for a tree in the m^{th} block will be

$$(\rho \cdot P)_m = \sum_{b=1}^{p} \left(\frac{1}{p^{b+1}}\right)_{|c(pm)_b|} \cdot c(pm)_b \tag{6.45}$$

The product in (6.45) simplifies to $1/p^{m+1}$ (Appendix B.1). We have that $(\rho.P)_m = (\rho)_m$ for all $m \in [1, p]$ which implies that $\rho.P = \rho$. It is also obvious that the elements of ρ sum up to 1 since there are p blocks with trees in each block summing up to $1/p$. Hence, ρ is the stationary distribution for the Markov chain of depth 2 tree shapes. □

Let us calculate the average cost under the policy π. The expected cost of a tree $t = (\mu, \{t_{0,j_1}, \cdots, t_{0,j_m}\})$ is

$$E[c(t, \pi(t))] = \frac{m-1}{m} \tag{6.46}$$

Therefore, the average cost is (split into blocks):

$$\begin{aligned}
&\sum_{t\in p1} \frac{1}{p^2} E[c(t,\pi(t))] + \sum_{t\in p2} \frac{1}{p^3} E[c(t,\pi(t))] + \cdots + \sum_{t\in pp} \frac{1}{p^{p+1}} E[c(t,\pi(t))] \\
&= \sum_{t\in p1} \frac{1}{p^2}\frac{1-1}{1} + \sum_{t\in p2} \frac{1}{p^3}\frac{2-1}{2} + \cdots + \sum_{t\in pp} \frac{1}{p^{p+1}}\frac{p-1}{p} \\
&= \frac{1}{p^2}\frac{1-1}{1}\cdot p + \frac{1}{p^3}\frac{2-1}{2}\cdot p^2 + \cdots + \frac{1}{p^{p+1}}\frac{p-1}{p}\cdot p^p \\
&= \sum_{m=1}^{p} \frac{1}{p}\frac{m-1}{m} \\
&= \sum_{m=1}^{p} \frac{1}{p}\left(1 - \frac{1}{m}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{p}p - \frac{1}{p}\sum_{m=1}^{p}\frac{1}{m} \\
&= 1 - \frac{\mathbb{H}_p}{p}.
\end{aligned}
\tag{6.47}
$$

In the finite horizon criteria, we had the following expression for the value function under the greedy policy π.

$$
W_\pi(n,t) = n - \frac{1}{m} - \frac{n-2}{p}\cdot \mathbb{H}_p - \frac{1}{m}\sum_{r=1}^{m}\frac{1}{j_r} \text{ for } 2 \le n \le N. \tag{6.48}
$$

The average cost of this policy in the infinite horizon is:

$$
\begin{aligned}
&\lim_{N\to\infty}\frac{W_\pi(N,t)}{N} \\
&= \lim_{N\to\infty}\frac{1}{N}\left(N - \frac{1}{m} - \frac{N-2}{p}\cdot \mathbb{H}_p - \frac{1}{m}\sum_{r=1}^{m}\frac{1}{j_r}\right) \\
&= \lim_{N\to\infty}\left(1 - \frac{1}{Nm} - \frac{N-2}{Np}\cdot \mathbb{H}_p - \frac{1}{Nm}\sum_{r=1}^{m}\frac{1}{j_r}\right) \\
&= 1 - \frac{\mathbb{H}_p}{p} + \lim_{N\to\infty}\left(-\frac{1}{Nm} - \frac{2}{Np}\cdot \mathbb{H}_p - \frac{1}{Nm}\sum_{r=1}^{m}\frac{1}{j_r}\right) \\
&= 1 - \frac{\mathbb{H}_p}{p}
\end{aligned}
\tag{6.49}
$$

This matches with (6.47). Now let us explore this Markov chain of tree shapes and analyse its properties. In the this section, we noticed that the transitions of states was invariant of the markings. Let us try to calculate the average number of nodes, average number of nodes created at each time step, and average number of nodes deleted at each time step.

6.4 Average number of nodes

For depth 2 trees, we shall first calculate the total number of nodes of all trees :

$$
\sum_{m=1}^{p}\sum_{1\le j_1,\cdots,j_m\le p}\left((m+1)+(j_1+\cdots+j_m)\right) \tag{6.50}
$$

$$
\begin{aligned}
&= \sum_{m=1}^{p} p^m(m+1) + \sum_{m=1}^{p}\sum_{j_1=1}^{p}\cdots\sum_{j_m=1}^{p}(j_1+\cdots+j_m) \\
&= \sum_{m=1}^{p} p^m(m+1) + \sum_{m=1}^{p}\frac{p(p+1)}{2}\cdot p^{m-1}\cdot m \\
&= \sum_{m=1}^{p} p^m\left(m+1+\frac{p+1}{2}\cdot m\right) \qquad (6.51)
\end{aligned}
$$

In (6.50), the $(m+1)$ is from one root node and m sons. Each of the m sons have $j_1, \cdots, j_m$ sons/leaves of their own.

The total number of depth 2 trees is $p + p^2 + \cdots + p^p$. Hence, dividing (6.51) by the total number of depth 2 trees, we obtain the following simplified expression :

$$
\frac{5-p+p^p(p^3+2p^2-2p-5)}{2(p-1)(p^p-1)} \qquad (6.52)
$$

for $p > 1$. The formula in (6.52) for $p = 2, 3$ gives the values $31/6, 115/13$ respectively.

The above approach in (6.52) assumed uniform stationary probability of trees. However this is not the case as we know the stationary probabilities are not uniform. Hence using (6.51), we multiply the stationary probability of trees in each block to obtain :

$$
\sum_{m=1}^{p}\frac{1}{p^{m+1}}p^m\left(m+1+\frac{p+1}{2}\cdot m\right) = \sum_{m=1}^{p}\frac{1}{p}\left(m+1+\frac{p+1}{2}\cdot m\right) = \frac{7+4p+p^2}{4} \qquad (6.53)
$$

Using (6.53) for $p = 1, 2, 3$ we obtain $3, 19/4, 7$ respectively.

6.5 Average number of nodes created

Given a depth two tree $t = (\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\})$, let us proceed as below,

- Pick an $r \in [1, m]$ with probability $1/m$.
- To each of these j_r leaves, add between $1, \cdots, p$ new leaves with uniform probability $1/p$.

Therefore, for the tree t, the expected number of nodes added is :

$$\sum_{r=1}^{m} \frac{1}{m}\left(\frac{p+1}{2}\right) \cdot j_r \tag{6.54}$$

This is because adding new nodes to each of the j_r leaves are independent. Let Y_1 be the random variable that denotes the number of nodes added to the first leaf of the tree t_{i_r,j_r}, Y_2 denotes the number of nodes added to the second leaf of the tree t_{i_r,j_r}, and Y_{j_r} denotes the number of nodes added to the last leaf of the tree t_{i_r,j_r}. Then.

$$E[Y_1 + Y_2 + \cdots + Y_{j_r}] = E[Y_1] + E[Y_2] + \cdots + E[Y_{j_r}] = \left(\frac{p+1}{2}\right) \cdot j_r \tag{6.55}$$

And the probability of picking the subtree t_{i_r,j_r} is $1/m$. Therefore the expected number of nodes created for tree t is (6.54). The stationary probability of a tree with m sons is $1/p^{m+1}$. Hence summing over all possible j_r for all possible m, and multiplying it by the stationary probability, we obtain :

$$\sum_{m=1}^{p} \frac{1}{p^{m+1}} \sum_{1 \le j_1, \cdots, j_m \le p} \sum_{r=1}^{m} \frac{1}{m}\left(\frac{p+1}{2}\right) \cdot j_r \tag{6.56}$$

Example : For $p = 1, 2, 3, 4$, we obtain $1, 9/4, 4, 25/4$ respectively.

Equation (6.56) simplifies to :

$$\left(\frac{p+1}{2}\right)^2 \tag{6.57}$$

6.6 Average number of nodes deleted

Given a tree $t = (\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\})$, let us proceed as below.

- Pick an $r \in [1, m]$ with probability $1/m$.
- Remove the root node plus the sub-trees t_{i_q,j_q} for all $q \neq r$.

For a tree t, the expected number of nodes deleted is :

$$\sum_{r=1}^{m} \frac{1}{m}\left(1 + (m-1) + (j_1 + \cdots + j_m) - j_r\right) \tag{6.58}$$

Summing over all possible combinations of number of leaves for each son, and then summing over all possible number of sons, and multiplying this with the stationary probabilities we obtain (simplification in Appendix B.2),

$$\sum_{m=1}^{p} \frac{1}{p^{m+1}} \sum_{1 \leq j_1, \cdots, j_m \leq p} \sum_{r=1}^{m} \frac{1}{m} \left(1 + (m-1) + (j_1 + \cdots + j_m) - j_r\right)$$
$$= \left(\frac{p+1}{2}\right)^2. \tag{6.59}$$

This is equal to the average number of nodes created in (6.57) as expected. If they were unequal, we would have increasing or decreasing sizes of trees as the process goes ahead in time.

6.7 Specific Parameters

Due to various complexities arising from the size of the state space in the MDP of budget 2 depth 2 trees, we shall focus on some specific policies and try to compute the optimal policy given certain parameter values.

6.7.1 Policy of Marking Sons Only

We compute the average cost of the Greedy depth 1 policy for depth 2 trees with budget k. Since no leaf is marked, the trees in U will not have any marked sons. The expected cost of a tree $t = (\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\})$ will be :

$$E[c(t, \pi(t))] = \frac{(m-k)^+}{m} \tag{6.60}$$

We know that the stationary distribution of tree shapes at depth 2 is $1/p^{m+1}$ for a tree in block m. Hence, the product of the distribution and the expected costs summed over all blocks and trees would be :

$$\frac{1}{p} \sum_{m=1}^{p} \frac{(m-k)^+}{m} = \frac{1}{p} \sum_{m=k+1}^{p} \frac{m-k}{m} = \frac{1}{p} \left(p - k - k(H_p - H_k)\right). \tag{6.61}$$

Let us have the condition that $k < p$, since otherwise the MDP will be zero cost always. We have listed some values of the average cost for a few combinations of p, k in Table 6.1.

Table 6.1: Table of p, k, and average costs

p	k	Average Cost
2.0	1.0	0.25
2.0	2.0	0.0
3.0	1.0	0.3889
3.0	2.0	0.1110
3.0	3.0	0.0
4.0	1.0	0.4791
4.0	2.0	0.2083
4.0	3.0	0.0625
4.0	4.0	0.0

p	k	Average Cost
20.0	1.0	0.8201
20.0	2.0	0.6902
20.0	3.0	0.5853
20.0	18.0	0.0076
20.0	19.0	0.0025
20.0	20.0	0.0
105.0	2.0	0.9097
105.0	3.0	0.8742
105.0	100.0	0.0013

There seems to be a gradual fall off in costs as k increases.

Let us try to solve the equation average costs$(p, k) = \alpha$, where α is some value in $[0, 1]$, for a fixed p and a variable k. This is an attempt to find out how much budget is needed for some pre-determined cost.

$$\alpha = \frac{1}{p}\left(p - k - k(\mathbb{H}_p - \mathbb{H}_k)\right) \tag{6.62}$$

$$p(1-\alpha) = k(1 + \mathbb{H}_p - \mathbb{H}_k). \tag{6.63}$$

We use the approximation for large k and p ; $\mathbb{H}_k \approx \ln k + \gamma$ and a similar expression for $\mathbb{H}_p$, where γ is the Euler constant (≈ 0.5772).

$$p(1-\alpha) \approx k\left(1 + \log\frac{p}{k}\right)$$
$$\frac{p}{k}(1-\alpha) \approx 1 + \log\frac{p}{k}$$

Using `Mathematica` we obtain :

$$k \approx \frac{(\alpha-1)p}{ProductLog\left[\frac{\alpha-1}{e}\right]} \tag{6.64}$$

where $ProductLog(z)$ is the value x that solves the equation $z = x \cdot e^x$

We examine how k varies as α increase in the interval $[0, 1]$ for two values of $p = 500, 2000$ in Figure 6.1 and Figure 6.2 respectively.

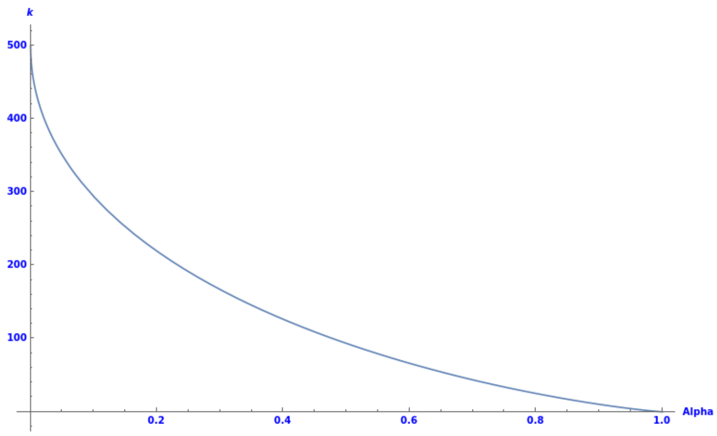

Figure 6.1: k versus α for $p = 500$

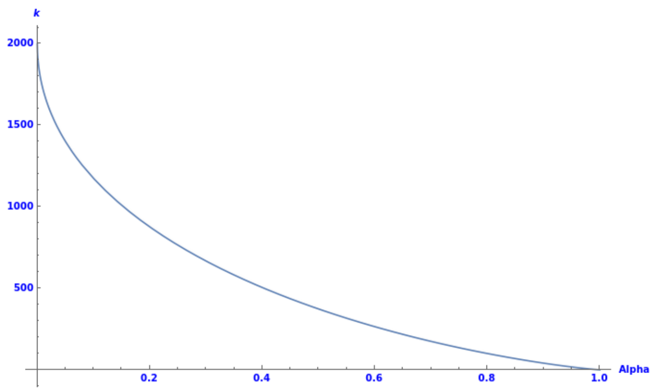

Figure 6.2: k versus α for $p = 2000$

As expected in both Figure 6.1 and Figure 6.2, α is closer to 0 as k is closer to p, and α increases to 1 as k decreases towards 0.

6.7.2 MDP with $p = 3, d = 2, k = 2$

For the MDP $p = 3, d = 2, k = 2$, we checked for the policy of marking the unmarked sons first. If there is budget left after marking an unmarked son or if all sons are already marked, then mark the first leaf (from left) of the first son. For Type $2b$ trees ($\{\mu, \{t_{1,j_1}, t_{1,j_2}\}\}$), mark the first two leaves of t_{1,j_1} or if $j_1 = 1$, mark that leaf and the first leaf of t_{1,j_2}. For the special case of tree $\{\mu, \{t_{1,1}\}\}$, just mark the leaf.

Using the same transition matrix of our numerical program for the Greedy Depth 1 policy on depth 2 trees, we obtain the average cost for $p = 3$ to be 0.1111 and for $p = 4$ to be 0.2083. This is consistent with the formula in (6.62),

$$\frac{1}{3}(3 - 2 - 2(\mathbb{H}_3 - \mathbb{H}_2)) = \frac{1}{3}\left(1 - 2\left(\frac{11}{6} - \frac{3}{2}\right)\right) = \frac{1}{9}$$
$$\frac{1}{4}(4 - 2 - 2(\mathbb{H}_4 - \mathbb{H}_2)) = \frac{1}{4}\left(2 - 2\left(\frac{25}{12} - \frac{3}{2}\right)\right) = \frac{5}{24}$$

Let us check for different greedy policies. We can mark the leaf of the leftmost subtree (Greedy Left), largest subtree (Greedy Large), or the smallest subtree (Greedy Small). The average cost under these variants of the greedy policy are :

Table 6.2: Average Cost of different Greedy policies

Policy/p	3	4
Greedy Depth 1	0.111111	0.208333
Greedy Small	0.067912	0.161568
Greedy Left	0.062802	0.160227
Greedy Large	0.054369	0.156907

As expected, in Table 6.2, the average cost under the policy Greedy Large is the least, followed by Greedy Left, and then Greedy Small. To get an idea of the enormous size of the Usable state space, Table 6.3 lists a few values of $|U|$ for different p.

Table 6.3: Table of Usable state space sizes for $k = 2, d = 2$

p value	Size of Usable state space - $\lvert U \rvert$
3	231
4	3,336
5	57,860
6	1,166,772
7	26,768,959

The complexity of analysis in the infinite horizon criterion has reached a peak here as the size of $|U|$ for $p = 7$ is in the tens of millions already! Even numerical programs are inefficient at computing the optimal policy. We truncate our analysis of the infinite horizon prefetching MDP for this thesis here.

Chapter 7

Conclusion

We analytically solved for the optimal policy for depth 1 trees with any budget and depth 2 trees with budget 1, for the finite horizon expected costs criterion and the infinite horizon average costs criterion. The optimal policy for these cases was the greedy policy of marking sons first. Moving to budget two complicated the analytical study. The complexity arose from the wide range of value functions that exist for different tree types in the finite horizon case, and the wide range of possible markings for the infinite horizon case. There was an interesting phenomena at the $n = 3$ horizon of the finite horizon MDP for particular kind of Type $2b$ trees where all three possible strategies were optimal for some range of p and size of subtrees. In the finite horizon study, we have structured the state space in a budget and policy dependent manner which would ease the search for an optimal policy given a particular tree.

The natural way of extending this structure to depth 2 trees with any budget k would be to alter the Tree type definition for $t = (\mu, \{t_{i_1,j_1}, \cdots, t_{i_m,j_m}\})$ as follows:

- Type 1 : $m - \sum_{r=1}^{m} i_r - k \geq 0$.
- Type $2a_1$: $m - \sum_{r=1}^{m} i_r = k - 1$.
- Type $2a_2$: $m - \sum_{r=1}^{m} i_r = k - 2$ and so on.

- For a general $1 \leq z \leq k-1$, Type $2a_z$: $m - \sum_{r=1}^{m} i_r = k - z$.
- Type 2b : $m - \sum_{r=1}^{m} i_r = 0$, $m > 1$.
- Type 2c : Simply the tree $(\mu, \{t_{1,j_1}\})$ where $1 \leq j_1 \leq p$. This is the trivial tree with one marked son.

One may ask why we are centered around the greedy policy? From the analysis for Type 1 and Type 2 trees, the other actions where one does not prioritize marking sons before leaves involve an extra cost of anywhere between $1/m$ to k/m. Our analysis has shown that for Type 1 and Type 2a trees, marking leaves that would yield lower future costs does not subdue the higher immediate cost of not marking sons. It might be reasonable to extend this to an arbitrary budget k.
For greater depths, we propose a certain way of partitioning the state space which might prove useful in solving for the optimal policy. Depth 2 trees involved the possibility of planning only one step in the future which was not better than focusing on the present. However, in depth 3 trees, the controller can plan two steps into the future which might be better than focusing on the present.
Perhaps, we could divide the depth d tree into $(d-1)$ depth two trees (overlapping depths), and assign equal fractions of the budget to each depth 2 tree. Then we could structure the state space as Type 1, Type $2a_z$, Type $3a_z$, $\cdots$, Type $(d-1)a_z$, Type $2b$, Type $3b$, $\cdots$,Type $(d-1)b$, Type $2c$, Type $3c$, $\cdots$, Type $(d-1)c$. Like before, $z \in [1, k-1]$. This structuring would entail solving for the optimal policy for each depth-2 tree to arrive at an optimal policy for the depth d tree. We are yet to explore this prospect.

The analysis of the Markov chain of depth two tree shapes provided some insight for the policies to be evaluated using the stationary probability approach in the infinite horizon criterion. When we evaluated the average cost for the MDP with $p = 3, k = 2, d = 2$, we obtained that a specific greedy policy of marking the leaf of the largest subtree (when all sons are marked) is optimal. This matches with the computation of the optimal policy in the finite horizon case for the same parameters. We faced complexity issues while extending to larger depths or larger budgets merely due to the enormous size of the state space. Moreover, numerical programs also seemed to be inefficient in calculating the optimal policy for large state spaces. It may prove useful to restrict ourselves to smaller and more realistic state spaces. The Markov Decision Process framework has been a useful tool in analysing the process of prefetching. Further work that builds on this thesis could be to relax one of the assumptions in our design of the MDP. Perhaps the surfer is allowed to move backwards or move with a non-uniform probability to one of the sons. These variants are interesting as they model a surfer's movement on a website more realistically. Analysis of such variants however would involve a great deal of complexity as the simpler MDP studied in this thesis proved to be challenging.

Bibliography

[1] [*MP94*] Martin L. Puterman, Markov Decision Processes, 1994, Wiley-Interscience.

[2] [*SR83*] Sheldon Ross, Introduction to Stochastic Dynamic Programming, 1983, Academic Press.

Appendix A

Depth 2 Trees

A.1 Budget 1 recurrence

We evaluate the value function under policy π at $n = 2$,

$$
\begin{aligned}
W_\pi(2,t) &= c(t,\pi(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t,\pi(t),t') \cdot W_\pi(1,t'), \\
&= c(t,\pi(t)) + \sum_{r=1}^{m} \sum_{t' \in \mathcal{D}(t_{i_r,j_r})} P(t,\pi(t),t') \cdot W_\pi(1,t') \\
&= c(t,\pi(t)) + \sum_{r=1}^{m} \frac{1}{mp^{j_r}} \sum_{1 \le f_1, f_2, \cdots, f_{j_r} \le p} \left(1 - \frac{1}{j_r}\right) \qquad \text{(A.1)} \\
&= c(t,\pi(t)) + \sum_{r=1}^{m} \frac{1}{mp^{j_r}} \cdot \left(p^{j_r} - \frac{p^{j_r}}{j_r}\right) \\
&= c(t,\pi(t)) + \sum_{r=1}^{m} \left(\frac{1}{m} - \frac{1}{m \cdot j_r}\right) \\
&= c(t,\pi(t)) + 1 - \frac{1}{m} \sum_{r=1}^{m} \frac{1}{j_r} \\
&= 1 - \frac{1}{m} + 1 - \frac{1}{m} \sum_{r=1}^{m} \frac{1}{j_r}
\end{aligned}
$$

$$= 2 - \frac{1}{m} - \frac{1}{m}\sum_{r=1}^{m}\frac{1}{j_r}. \tag{A.2}$$

Note that in (A.1), the innermost summation is independent of f_q for all $q \in \{1, \cdots, j_r\}$.

$$\begin{aligned}
W_\pi(3,t) &= c(t,\pi(t)) + \sum_{t' \in \mathcal{SD}(t)} P(t,\pi(t),t') \cdot W_\pi(2,t') \\
&= c(t,\pi(t)) + \sum_{r=1}^{m}\left(\frac{1}{mp^{j_r}} \sum_{1 \le f_1, f_2, \cdots, f_{j_r} \le p} W_\pi(2,(i_r,\{t_{0,f_1},\cdots,t_{0,f_{j_r}}\}))\right) \\
&= c(t,\pi(t)) + \sum_{r=1}^{m}\left(\frac{1}{mp^{j_r}} \sum_{1 \le f_1, f_2, \cdots, f_{j_r} \le p} \left(2 - \frac{1}{j_r} - \frac{1}{j_r}\cdot\sum_{q=1}^{j_r}\frac{1}{f_q}\right)\right) \quad \text{(A.3)} \\
&= c(t,\pi(t)) + 2 - \frac{1}{m}\cdot\sum_{r=1}^{m}\frac{1}{j_r} - \sum_{r=1}^{m}\left(\frac{1}{mp^{j_r}j_r}\cdot\sum_{1 \le f_1, f_2, \cdots, f_{j_r} \le p}\sum_{q=1}^{j_r}\frac{1}{f_q}\right) \\
&= 3 - \frac{1}{m} - \frac{1}{m}\cdot\sum_{r=1}^{m}\frac{1}{j_r} - \frac{1}{m}\cdot\sum_{r=1}^{m}\left(\frac{1}{p^{j_r}j_r}\cdot\sum_{1 \le f_1, f_2, \cdots, f_{j_r} \le p}\sum_{q=1}^{j_r}\frac{1}{f_q}\right) \\
&= 3 - \frac{1}{m} - \frac{1}{m}\cdot\sum_{r=1}^{m}\left(\frac{1}{j_r} + \mathbb{M}(1,j_r)\right), \quad \text{(A.4)}
\end{aligned}$$

where we have used the value of $W_\pi(2,t)$ from (A.2).

A.2 $\mathbb{M}(n,j)$ recurrence

Recall the summation of harmonic series as :

$$\mathbb{H}_p = \sum_{f=1}^{p}\frac{1}{f}. \tag{A.5}$$

Then,

$$\mathbb{M}(1,1) = \frac{1}{p}\sum_{1 \le f_1 \le p}\frac{1}{f_1} = \frac{1}{p}\cdot\mathbb{H}_p \tag{A.6}$$

$$
\begin{aligned}
\mathbb{M}(1,2) = \frac{1}{2p^2} \sum_{1 \le f_1, f_2 \le p} \left(\frac{1}{f_1} + \frac{1}{f_2} \right) &= \frac{1}{2p^2} \sum_{f_1=1}^{p} \sum_{f_2=1}^{p} \left(\frac{1}{f_1} + \frac{1}{f_2} \right) \\
&= \frac{1}{2p^2} \cdot \left(\sum_{f_1=1}^{p} \sum_{f_2=1}^{p} \frac{1}{f_1} + \sum_{f_1=1}^{p} \sum_{f_2=1}^{p} \frac{1}{f_2} \right) \\
&= \frac{1}{2p^2} \cdot (p \cdot \mathbb{H}_p + p \cdot \mathbb{H}_p) \\
&= \frac{1}{2p^2} \cdot (2p \cdot \mathbb{H}_p) \\
&= \frac{1}{p} \cdot \mathbb{H}_p.
\end{aligned}
$$

Hence, for any $2 \le j \le p$:

$$
\begin{aligned}
\mathbb{M}(1,j) &= \frac{1}{jp^j} \sum_{1 \le f_1, f_2, \cdots, f_j \le p} \left(\frac{1}{f_1} + \frac{1}{f_2} + \cdots + \frac{1}{f_j} \right) \\
&= \frac{1}{jp^j} \sum_{f_1=1}^{p} \sum_{f_2=1}^{p} \cdots \sum_{f_j=1}^{p} \left(\frac{1}{f_1} + \frac{1}{f_2} + \cdots + \frac{1}{f_j} \right) \\
&= \frac{1}{jp^j} \cdot \left(\sum_{f_1=1}^{p} \cdots \sum_{f_j=1}^{p} \frac{1}{f_1} + \cdots + \sum_{f_1=1}^{p} \cdots \sum_{f_j=1}^{p} \frac{1}{f_j} \right) \\
&= \frac{1}{jp^j} \cdot \left(p^{j-1} \cdot \mathbb{H}_p + \cdots + p^{j-1} \cdot \mathbb{H}_p \right) \\
&= \frac{1}{jp^j} \cdot (j \cdot p^{j-1} \cdot \mathbb{H}_p) \\
&= \frac{1}{p} \cdot \mathbb{H}_p. \qquad \text{(A.7)}
\end{aligned}
$$

We have a simplified expression for $\mathbb{M}(1,j)$. Now, let us simplify $\mathbb{M}(2,j)$,

$$
\begin{aligned}
\mathbb{M}(2,j) &= \frac{1}{jp^j} \cdot \sum_{1 \le f_1, f_2, \cdots, f_j \le p} \sum_{q=1}^{j} \left(\frac{1}{f_q} + \mathbb{M}(1, f_q) \right) \\
&= \frac{1}{jp^j} \cdot \left(\sum_{f_1=1}^{p} \cdots \sum_{f_j=1}^{p} \sum_{q=1}^{j} \frac{1}{f_q} + \sum_{f_1=1}^{p} \cdots \sum_{f_j=1}^{p} \sum_{q=1}^{j} \frac{1}{p} \cdot \mathbb{H}_p \right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{jp^j} \cdot \left(p^{j-1} \cdot j \cdot \mathbb{H}_p + \frac{1}{p} \cdot \mathbb{H}_p \cdot p^j \cdot j \right) \\
&= \frac{1}{p} \cdot \mathbb{H}_p + \frac{1}{p} \cdot \mathbb{H}_p \\
&= \frac{2}{p} \cdot \mathbb{H}_p.
\end{aligned}
$$

Thus by a simple induction, we obtain for $2 \leq n \leq N$,

$$
\begin{aligned}
\mathbb{M}(n,j) &= \frac{1}{jp^j} \cdot \sum_{1 \leq f_1, f_2, \cdots, f_j \leq p} \sum_{q=1}^{j} \left(\frac{1}{f_q} + \mathbb{M}(n, f_q) \right) \\
&= \frac{1}{jp^j} \cdot \left(\sum_{f_1=1}^{p} \cdots \sum_{f_j=1}^{p} \sum_{q=1}^{j} \frac{1}{f_q} + \sum_{f_1=1}^{p} \cdots \sum_{f_j=1}^{p} \sum_{q=1}^{j} (n-1) \cdot \frac{1}{p} \cdot \mathbb{H}_p \right) \\
&= \frac{1}{jp^j} \cdot \left(p^{j-1} \cdot j \cdot \mathbb{H}_p + \frac{n-1}{p} \cdot \mathbb{H}_p \cdot p^j \cdot j \right) \\
&= \frac{1}{p} \cdot \mathbb{H}_p + (n-1) \cdot \frac{1}{p} \cdot \mathbb{H}_p \\
&= \frac{n}{p} \cdot \mathbb{H}_p.
\end{aligned}
$$

A.3 $Tr(2:1,1)$

We compute a more explicit expression for $Tr(2:1,1)$.

$$
\begin{aligned}
2Tr(2:1,1) = &\sum_{f_2=1}^{p} \frac{(f_2-4)^+}{f_2} \; [f_1 = 1] \\
&+ \sum_{f_2=1}^{p} \frac{(f_2-4)^+}{f_2} \; [f_1 = 2] \\
&+ \sum_{f_2=1}^{p} + \frac{(f_3-2)^+}{f_2} \; [f_1 = 3] \\
&+ \sum_{f_2=1}^{2} \frac{(f_1-4)^+}{f_1} + \frac{(f_1-3)^+}{f_1} + (\sum_{f_2=4}^{4} \frac{(f_1-2)^+}{f_1} + \frac{(f_2-4)^+}{f_2}) \\
&+ \sum_{f_2=5}^{p} \left(\frac{(f_1-4)^+}{f_1} + \frac{(f_2-2)^+}{f_2} \right) \; [f_1 = 4]
\end{aligned}
$$

$$+\sum_{f_2=1}^{2}\frac{(f_1-4)^+}{f_1}+\frac{(f_1-3)^+}{f_1}+\sum_{f_2=4}^{5}\left(\frac{(f_1-2)^+}{f_1}+\frac{(f_2-4)^+}{f_2}\right)$$
$$+\sum_{f_2=6}^{p}\left(\frac{(f_1-4)^+}{f_1}+\frac{(f_2-2)^+}{f_2}\right)\ [f_1=5]$$
$$\ldots$$
$$+\sum_{f_2=1}^{2}\frac{(f_1-4)^+}{f_1}+\frac{(f_1-3)^+}{f_1}+\sum_{f_2=4}^{p}\left(\frac{(f_1-2)^+}{f_1}+\frac{(f_2-4)^+}{f_2}\right)$$
$$[f_1=p].$$

Clubbing similar terms together we obtain :

$$\begin{aligned}2Tr(2:1,1)=4\mathbb{H}_{p4}+2\mathbb{H}_{p3}+\sum_{f_1=4}^{p}\sum_{f_2=4}^{f_1}\left(\frac{(f_1-2)^+}{f_1}+\frac{(f_2-4)^+}{f_2}\right)\\ +\sum_{f_1=4}^{p-1}\sum_{f_2=f_1+1}^{p}\left(\frac{(f_1-4)^+}{f_1}+\frac{(f_2-2)^+}{f_2}\right).\end{aligned} \tag{A.8}$$

We simplify (A.8) using `Mathematica` and obtain :

$$Tr(2:1,1)=p^2+\frac{19p}{3}+5-(4p+6)\mathbb{H}_p. \tag{A.9}$$

A.4 $Tr(3:1,1,0)$

We shall split the summation over f_1 into three parts. The first part is for $f_1=1,2$, the second for $f_1=3$, and the third for $f_1\in[4,p]$. (A.10) details the contribution of $f_1=1,2$ to $3Tr(3:1,1,0)$.

$$\begin{aligned}2\Bigg\{\sum_{f_2=1}^{2}\sum_{f_3=4}^{p}\frac{f_3-3}{f_3}+\sum_{f_3=3}^{p}\frac{f_3-2}{f_3}\\ +\sum_{f_2=4}^{p}\Bigg\{\frac{2(f_2-3)}{f_2}+\frac{f_2-2}{f_2}+\sum_{f_3=4}^{f_2}\left(\frac{f_3-3}{f_3}+\frac{f_2-2}{f_2}\right)\\ +\sum_{f_3=f_2+1}^{p}\left(\frac{f_2-3}{f_2}+\frac{f_3-2}{f_3}\right)\Bigg\}\Bigg\}.\end{aligned} \tag{A.10}$$

Similarly for $f_1 = 3$, we have (A.11) as the contribution to $3Tr(3:1,1,0)$

$$\begin{aligned} &\sum_{f_2=1}^{2}\left(\frac{1}{3}+\sum_{f_3=4}^{p}\frac{f_3-2}{f_3}\right)+\sum_{f_3=1}^{2}\frac{1}{3}+\sum_{f_3=3}^{p}\left(\frac{f_3-2}{f_3}+\frac{1}{3}\right) \\ &+\sum_{f_2=4}^{p}\left\{\frac{2(f_2-2)}{f_2}+\frac{f_2-2}{f_2}+\frac{1}{3}+\sum_{f_3=4}^{p}\left(\frac{f_3-2}{f_3}+\frac{f_2-2}{f_2}\right)\right\}. \end{aligned} \tag{A.11}$$

We shall now detail the contribution from $f_1 \in [4,p]$ to $3Tr(3:1,10)$ in (A.12).

$$\begin{aligned} &\sum_{f_1=4}^{p}\left\{\sum_{f_2=1}^{2}\left(\frac{2(f_1-3)}{f_1}+\frac{f_1-2}{f_1}+\sum_{f_3=4}^{f_1}\left(\frac{f_3-3}{f_3}+\frac{f_1-2}{f_1}\right)\right.\right. \\ &+\left.\sum_{f_3=f_1+1}^{p}\left(\frac{f_3-2}{f_3}+\frac{f_1-3}{f_1}\right)\right)+\sum_{f_3=1}^{2}\frac{f_1-2}{f_1}+\sum_{f_3=3}^{p}\left(\frac{f_3-2}{f_3}+\frac{f_1-2}{f_1}\right) \\ &+\sum_{f_2=4}^{f_1}\left(2\left(\frac{f_2-3}{f_2}+\frac{f_1-2}{f_1}\right)+\frac{f_2-2}{f_2}+\frac{f_1-2}{f_1}\right. \\ &+\sum_{f_3=4}^{f_2}\left(\frac{f_3-3}{f_3}+\frac{f_2-2}{f_2}+\frac{f_1-2}{f_1}\right) \\ &+\left.\sum_{f_3=f_2+1}^{p}\left(\frac{f_3-2}{f_3}+\frac{f_2-3}{f_2}+\frac{f_1-2}{f_1}\right)\right) \\ &+\sum_{f_2=f_1+1}^{p}\left(2\left(\frac{f_2-2}{f_2}+\frac{f_1-3}{f_1}\right)+\frac{f_2-2}{f_2}+\frac{f_1-2}{f_1}\right. \\ &+\sum_{f_3=4}^{f_1}\left(\frac{f_3-3}{f_3}+\frac{f_2-2}{f_2}+\frac{f_1-2}{f_1}\right) \\ &+\left.\left.\sum_{f_3=f_1+1}^{p}\left(\frac{f_3-2}{f_3}+\frac{f_2-2}{f_2}+\frac{f_1-3}{f_1}\right)\right)\right\} \end{aligned} \tag{A.12}$$

Adding (A.10), (A.11), and (A.12) we obtain the value of $3Tr(3:1,1,0)$. We simplified it using `Mathematica` to obtain :

$$Tr(3:1,1,0)=\frac{1}{3}\left(\frac{15}{2}+3p(8+p(4+p))-(19+3p(5+3p))\mathbb{H}_p\right) \tag{A.13}$$

A.5 Type 1 trees at $n = 3$

We detail how we obtain the $\mathbb{H}_{p2}$ term.

$$
\begin{aligned}
&\sum_{r=1}^{m} \frac{1}{m \cdot p^{j_r}} \sum_{1 \leq f_1, \cdots, f_{j_r} \leq p} \frac{1}{j_r} \sum_{q=1}^{j_r} \frac{(f_q - 2)^+}{f_q} \\
&= \sum_{r=1}^{m} \frac{1}{m \cdot p^{j_r} \cdot j_r} \sum_{f_1=1}^{p} \cdots \sum_{f_{j_r}=1}^{p} \left(\frac{(f_1 - 2)^+}{f_1} + \cdots + \frac{(f_{j_r} - 2)^+}{f_{j_r}} \right) \\
&= \frac{1}{m} \sum_{r=1}^{m} \frac{1}{p^{j_r} j_r} \left(\sum_{f_1=1}^{p} \cdots \sum_{f_{j_r}=1}^{p} \frac{(f_1 - 2)^+}{f_1} + \cdots + \sum_{f_1=1}^{p} \cdots \sum_{f_{j_r}=1}^{p} \frac{(f_{j_r} - 2)^+}{f_{j_r}} \right) \\
&= \frac{1}{m} \sum_{r=1}^{m} \frac{1}{p^{j_r} j_r} \left(p^{j_r - 1} \cdot \mathbb{H}_{p2} + \cdots + p^{j_r - 1} \cdot \mathbb{H}_{p2} \right) \\
&= \frac{1}{m} \sum_{r=1}^{m} \frac{1}{p^{j_r} j_r} \cdot j_r \cdot p^{j_r - 1} \cdot \mathbb{H}_{p2} \\
&= \frac{1}{m} \sum_{r=1}^{m} \frac{1}{p} \mathbb{H}_{p2} \\
&= \mathbb{H}_{p2}.
\end{aligned}
\tag{A.14}
$$

Appendix B

Infinite Horizon Average Costs

B.1 Stationary Distribution for depth 2 trees

We compute the product of the stationary distribution and the transition matrix for a tree in the pm block.

$$
\begin{aligned}
(\rho \cdot P)_m &= \sum_{b=1}^{p} \frac{1}{p^{b+1}} \left(\frac{1}{p^m} + \sum_{r=1}^{b-1} \frac{b-r}{bp^m} \cdot \binom{b}{r} (p-1)^r \right) \\
&= \sum_{b=1}^{p} \frac{1}{p^{b+1}} \left(\sum_{r=0}^{b-1} \frac{b-r}{bp^m} \cdot \binom{b}{r} (p-1)^r \right) \\
&= \frac{1}{p^m} \sum_{b=1}^{p} \frac{1}{p^{b+1}} \left(\sum_{r=0}^{b-1} \frac{b-r}{b} \cdot \binom{b}{r} (p-1)^r \right) \\
&= \frac{1}{p^m} \sum_{b=1}^{p} \frac{1}{p^{b+1}} \left(\sum_{r=0}^{b-1} \frac{b-r}{b} \cdot \frac{b!}{(b-r)!r!} (p-1)^r \right) \\
&= \frac{1}{p^m} \sum_{b=1}^{p} \frac{1}{p^{b+1}} \left(\sum_{r=0}^{b-1} \frac{(b-1)!}{(b-1-r)!r!} \cdot (p-1)^r \right) \\
&= \frac{1}{p^m} \sum_{b=1}^{p} \frac{1}{p^{b+1}} \left(\sum_{r=0}^{b-1} \binom{b-1}{r} \cdot (p-1)^r \right) \\
&= \frac{1}{p^m} \sum_{b=1}^{p} \frac{1}{p^{b+1}} \cdot p^{b-1} \qquad \text{(B.1)}
\end{aligned}
$$

$$= \frac{1}{p^m} \cdot \frac{1}{p} = \frac{1}{p^{m+1}} \tag{B.2}$$

In (B.1), we have used the Binomial expansion theorem.

B.2 Average number of nodes deleted

The simplification for the average number of nodes deleted is as follows :

$$\begin{aligned}
&\sum_{m=1}^{p} \frac{1}{p^{m+1}} \sum_{1 \le j_1, \cdots, j_m \le p} \sum_{r=1}^{m} \frac{1}{m} \left(1 + (m-1) + (j_1 + \cdots + j_m) - j_r\right) \\
&= \sum_{m=1}^{p} \frac{1}{mp^{m+1}} \sum_{1 \le j_1, \cdots, j_m \le p} \sum_{r=1}^{m} \left(m + (j_1 + \cdots + j_m) - j_r\right) \\
&= \sum_{m=1}^{p} \frac{1}{mp^{m+1}} p^m \cdot m \cdot m + \sum_{m=1}^{p} \frac{1}{mp^{m+1}} \sum_{1 \le j_1, \cdots, j_m \le p} \sum_{r=1}^{m} (j_1 + \cdots + j_m - j_r) \\
&= \sum_{m=1}^{p} \frac{m}{p} + \frac{1}{mp^{m+1}} \sum_{j_1=1}^{p} \cdots \sum_{j_m=1}^{p} \sum_{r=1}^{m} (j_1 + \cdots + j_m - j_r) \qquad \text{(B.3)} \\
&= \frac{p+1}{2} + \sum_{m=1}^{p} \frac{1}{mp^{m+1}} \left(\frac{p(p+1)}{2} \cdot p^{m-1} \cdot (m-1) \cdot m \right) \\
&= \frac{p+1}{2} + \sum_{m=1}^{p} \frac{(p+1)(m-1)}{2p} \\
&= \frac{p+1}{2} + \frac{(p+1)}{2p} \sum_{m=1}^{p} (m-1) \\
&= \frac{p+1}{2} + \frac{(p+1)}{2p} \frac{p-1}{2} \\
&= \frac{2p + 2 + p^2 - 1}{4} \\
&= \left(\frac{p+1}{2} \right)^2 . \qquad \text{(B.4)}
\end{aligned}$$