Names and Matriculation numbers :
Amrita Bhattacherjee - 441338
Debabrata Ghosh - 441275
Ilya Fradlin - 441264

Probabilistic Programming
Exercise Sheet 3 Solutions
Group 7

---

1. (a) Check sheet03_ex1a.wppl for the code.

   (b) Check sheet03_ex1b.wppl for the code.

   (c) We can understand the impact of the parameter σ^2 of the proposal distribution through the samples accepted during our run of the algorithm.

   A really small variance of the proposal distribution leads to new samples which are almost always accepted as acceptance rate r is always close to 1 for continuous distributions since proposed moves are small. This issue can also be noticed with our proposed distribution in this problem.

   Also, since the movements in each step is small, the transitions are concentrated into a specific region leading to poor coverage of the distribution globally. So, the Markov chain might not cover the entire distribution space especially if we have multiple modes, in this situation the samples might converge to one of the modes and never move to other ones.

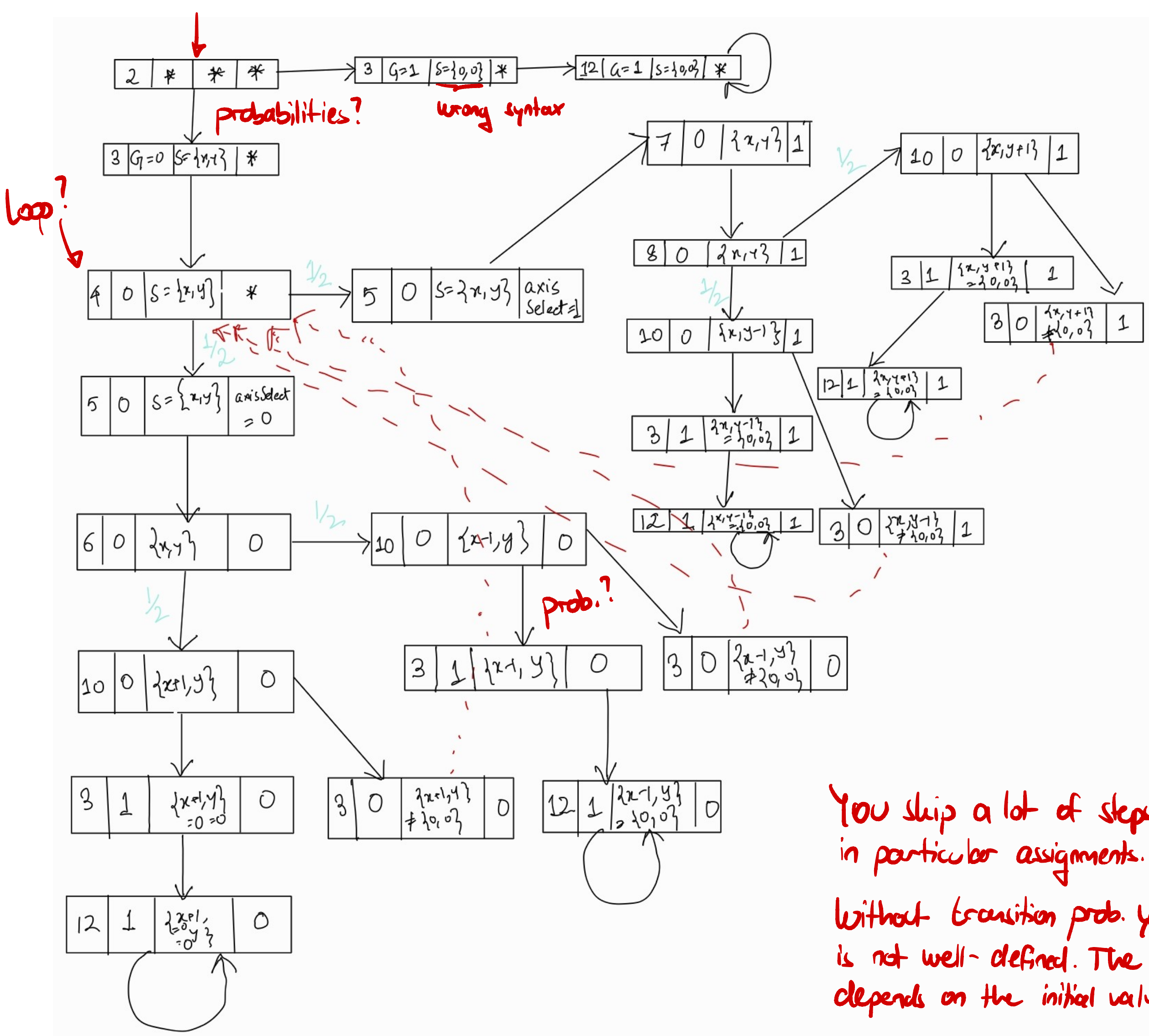
   On the other hand, if we used large variance in the proposal distribution then most transitions might have been proposed in regions of very little probability and many moves are rejected, leading to the ~~samples~~ stuck with a single value for many iterations. So, we need to find a middle ground. ✓

2. (a) The corresponding pGCL program for the stochastic process defined is as follows :

13/25

```
int TravelZ2(int x_0, y_0) {   Assign x0,y0 to variables
    bool G := (x_0 = 0) ∧ (y_0 = 0);
    while(¬G) {
        int axisSelect := 0 [0.5] axisSelect := 1;
        if(axisSelect=0){
            (x_0 := (x_0 + 1) [0.5] x_0 := (x_0 − 1));
        } else {
            (y_0 := (y_0 + 1) [0.5] y_0 := (y_0 − 1));
        }
        G := (x_0 = 0) ∧ (y_0 = 0);
    }
    return G;
}
```

(b) To get the operational semantics of the pGCL program defined in (a) as a Markov chain for one loop iteration we assume initial states $x_0 = xandy_0 = y$. See the figure below for the corresponding operational semantics:



You skip a lot of steps, in particular assignments.

Without transition prob. your MC is not well-defined. The execution depends on the initial values of x_0, y_0.

Sink states are missing.

Overall: -6P

3. 18/25

(a)

Markov Chain with rewards

We consider a six-sided fair die. We propose the reward Markov chain (D, r), where D is a Markov chain with state space $\Sigma = \{1,2,3,4,5,6\}$ and, $r : \Sigma \to \mathbb{R}$ is the reward function s.t.

$$r(\sigma) = 1 \quad \forall \sigma \in \Sigma$$

and goal state $G = \{6\}$ since we want to compute the expected # trials needed to see a six for the first time.

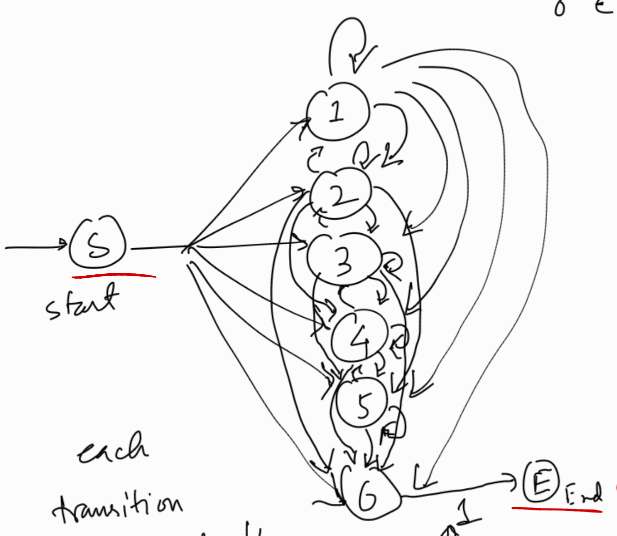
We know that expected rewards in finite Markov chain can be computed in polynomial time by solving a system of linear equations.

We also have a finite Markov chain in this case.

also $P(\sigma, \sigma') = 1/6 \quad \forall \sigma, \sigma' \in \Sigma$.

For $Pr(\sigma \models \Diamond G) = 1 \quad \forall \sigma \in \Sigma$, we take the variable $y_\sigma \in \mathbb{R}$ for $\sigma \in \Sigma$

$$y_\sigma = \begin{cases} 0 & \text{if } \sigma \in G \\ r(\sigma) + \sum_{\sigma' \in \Sigma} P(\sigma, \sigma') \cdot y_{\sigma'} & \text{o. w.} \end{cases}$$



each transition with prob 1/6 except

Missing In your definition, Missing ~~Probability~~ Probability mass. Reward?

Let N be the # trials until the MC hits 6 for the first time. If it hits 6 at the first step (with prob 1/6) we stop or with prob 5/6 it hits any other state we continue. So the expected # trials =

$$\frac{1}{6}\times 1 + 2\times\frac{5}{6}\times\frac{1}{6} + 3\times\frac{5}{6}\times\frac{5}{6}\times\frac{1}{6} + \cdots$$

$$= \frac{1}{6}\sum_{i\geq 1} i\left(\frac{5}{6}\right)^{i-1}$$

$$= \frac{1}{6}\sum_{k\geq 0}(k+1)\left(\frac{5}{6}\right)^{k}$$ (substituting $k = i-1$)

$$= \frac{1}{6}\sum_{k\geq 0} k\left(\frac{5}{6}\right)^{k} + \frac{1}{6}\sum_{k\geq 0}\left(\frac{5}{6}\right)^{k}$$

$$= \cancel{\frac{1}{6}}\times\frac{5/6}{\left(\cancel{1/6}\right)^2 \; 1/6} + \frac{1}{6}\cdot\frac{1}{1-5/6}$$

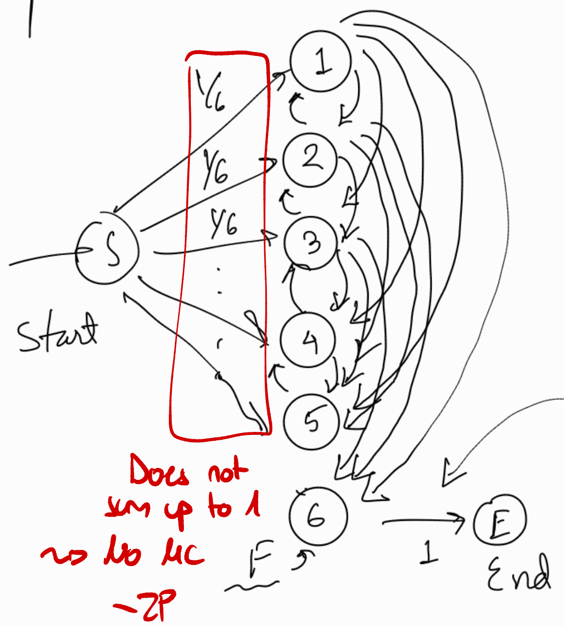
[By using the hint given and property of sum of Geometric Progression].

$$= 5 + \cancel{\frac{1}{6}}\times\frac{1}{\cancel{1/6}}$$

$$= 5 + 1$$

$$= \boxed{6}$$ ✓

3.
b) The only major condition changed in this case is that we can't go the forbidden state 6 in our first step. So the Markov chain changes accordingly (say D')



All transition probability is 1/6 except

If 6 is forbidden then you have prob. 0 to reach E when avoiding forbidden states: $Pr(1, \Diamond E \mid \neg \Diamond 6) = 0$ −5P

Since first trial was no six, the minimum # trials to get a six is 2.

So all the paths such that

$$\pi = \{\sigma_0 \dots \sigma_n \mid \sigma_0 = 6\} \notin Paths(D')$$

16/25

④ a) In order to prove that $(D \to D, \preceq)$ is a partial order we will show:

i) Reflexivity [redacted]:

let f be a monotonic function in D, $f \in D \to D$, ~~for~~ $\forall d \in D$ $f(d) \in D$, since $(D, \sqsubseteq)$ is a complete lattice [(thus a partial order)] we know $f(d) \sqsubseteq f(d)$ {from monotonicity of f $(D, \sqsubseteq)$}

reflexivity

−1P

$\Longleftrightarrow$ $f \preceq f$

by defenition

ii) transitivity:

let $f, g, h \in D \to D$ such that $f \preceq g$, $g \preceq h$, thus, by defenition, $\forall d \in D$ $f(d) \sqsubseteq g(d)$, $g(d) \sqsubseteq h(d)$. by transitivity of $(D, \sqsubseteq)$ we get $f(d) \sqsubseteq h(d)$ for $\forall d \in D$, thus $f \preceq \cancel{g}$. h

iii) antisymmetry:

let $f, g \in D \to D$, such that $f \preceq g$, $g \preceq f$. thus, by defenition: $\forall d \in D$ $f(d) \sqsubseteq g(d)$, $g(d) \sqsubseteq f(d)$. from antisymmetry of $(D, \sqsubseteq)$ we get $f(d) = g(d)$. since this is true ~~for~~ $\forall d \in D$ we can conclude $f = g$! ✓

b) we've seen that $(D \to D, \sqsubseteq)$ is indeed a partial order, let there B a subset $S \subseteq D \to D$ of monotonic functions in D.

we construct a function f, which for any given $d \in D$ $s(d) \sqsubseteq f(d)$ ~~for~~ $\forall s \in S$, such a function exist since we can just define: $f(d) := \max\{s(d) \mid s \in S\}$. such a function

Use $\sqcup$ instead of max. Why does this exist?

is also still monotonic since $\forall s \in S$ are monotonic as well $\Rightarrow$ if $d_1, d_2 \in D$, $d_1 \sqsubseteq d_2$

let $s_i(d_1)$ B the max value for d_1 out of S and $s_j(d_2)$ B the max value for d_2 out of S,

we get $s_j(d_1) \sqsubseteq s_i(d_1) \sqsubseteq s_i(d_1) \sqsubseteq s_j(d_2)$

why? is $s_j \leq s_i$? why?

from transitivity: $s_i(d_1) \sqsubseteq s_j(d_2) \Rightarrow s(d_1) \sqsubseteq f(d_2)$!

f is monotonic as well $\Rightarrow f \in D \to D$.

by definition f is an upper bound of S. ✓

now lets assume there exist g which is also an upper bound of S and $g \sqsubseteq f$, thus $\exists d \in D$

$\forall s_i \in S$ $s_i(d) \sqsubseteq g(d) \sqsubseteq f[d]$ contradicting the fact

one d or all? Why?

that $f(d)$ is defined as the max of $\{s(d) \mid s \in S\}$

$\Rightarrow f$ is lub \ supremum of S!

$\Rightarrow$ similiraly we can define function h by the minimum values and get that h is the infimum of S (✓)

Make sure that F, h depend on the set S

thus F, h are $\in D \to D$ and are the supremum and infimum of an arbitrary $S \subseteq D \to D$.

thus $(D \to D, \sqsubseteq)$ is a complete lattice.

Overall: -2P

$\square$

c) The statement is correct:

we'll prove it by using the hint.

i) first well prove that because D is satisfying the ACC, $D \to D$ also satisfies the same condition. let there be a function $f \in (D \to D)$

Your reasoning is not clear.

since D is ACC then f's Image and Pre Image is also limited! ?

Thus making a chain $F \leq (D \to D)$ follow the rules required from ACC.

Why does this imply ACC for $D \to D$? -3P

ii) In our scenario where D satisfies ACC lfp is actually a monotone function: $f_1, f_2 \in (D \to D)$ $f_1 \sqsubseteq f_2$, since D is limited (and basically finite) ?

D can still be infinite

? both functions will have $d_1, d_2 \in D$ where $d_1 = f_1(d_1)$ $d_2 = f_2(d_2)$ (from pigeonhole principle), and since we assumed $f_1 \sqsubseteq f_2$ and the function are monotonic we can derive that $lfp(f_1) \sqsubseteq lfp(f_2)$

Knaster-Tarski Theorem

You only argued for some fixed point, why also for the least? -3P

iii) $\Rightarrow$ Now we can use the hint:

$$\phi(\bigsqcup S) = \bigsqcup \phi(S)$$

$$\Downarrow$$

plugging In:

$$\phi = lfp$$

$$S = \mathcal{F} \leq (D \to D)$$

$$\Downarrow$$

$$lfp(\bigsqcup \mathcal{F}) = \bigsqcup lfp(f \mid f \in \mathcal{F})$$

(V)

Your reasoning is not clear in a lot of places. Please give more context/actual formal arguments when doing proofs.