The PIP model has 2 descriptions:

1. Local Continuous Time Markov Process

2. Global Poisson Process / Poisson Indel Process.

The notations are as follows:

Phylogeny is represented as **'T'.** The ancestor or root is denoted as **'*Ω*’.** The Insertion points are represented as Xi, were Xi is an element of **'T'.** Xi is any point on a branch of length, **b**. We say, **'Tx'.** is the rooted subtree of **'T'.** rooted at Xi, that is neglecting all the points in the TRUE tree that are not descendants of Xi. 'Σ' is set the set of characters/bases/nucleotides/nucleobases/AA.

Local Description: Insertion rate no longer depends on the sequence length, E.

How to determine next event?

1. We have **n** variables for substitutions (Which means all characters can undergo substitutions)

2. We have **n** variables for deletions, each of rate **‘μ’**

3. **1** variable for insertion with rate **'λ'**

Total 2n+1

At the i**nitial** state, the process initiated by sampling a number of characters at the root from a Poisson distribution with **λ** /**μ** which the asymptotic expected length. The sampling is done independently and identically according to substitution rate matrix, '**ϴ**'.

When an **insertion occurs:** Its position/location/site is selected uniformly at random. In detail, we assign each character a real value which lies in between 0 to 1. We assign real values to the characters in an increasing order fashion. When insertion happens, we sample a number which lies in between 0 to 1 and insert the character at the position which keeps the order unchanged (**this character is inserted at this position with a probability of 1 and of length 1**). In other words, the probability that evolutionary events occur at branching points is 0. Rate of this Poisson process has an atomic mass at the root of the tree.

In the next step we **visit each of the Insertion points**/locations one at a time. The order of visit is sampled randomly. When we visit this insertion location, 2 steps are carried out at this location. The first step prunes the tree from the insertion location, this subtree will then be used for the next stage. In the second step we simulate the character inserted with the subtree using **substitution-deletion CTMC**, which decide the fate of the inserted character. In other words, we can say that the simulation decides whether the inserted character survive to the leaf or gets deleted or substituted. The simulation yields the **homology path** which contains single-character history.

Each of the insertion location gives a column to the MSA. Each character in the column represents the characters at the leaves/taxon.

One of the properties of Poisson process is that the probability of insertion occurred on a branch of length will be proportional to branch length, b.