1 Previous Work

In previous work we have been considering the relationship between the Pure Parsimony and the Perfect Phylogeny Haplotyping. We have considered several g-lattice structures, from the Pure Parsimony Problem, to determine if they satisfy the requirement of supporting an unrooted perfect phylogeny, from the Perfect Phylogeny Haplotyping Problem. The g-lattices structures we considered were chains and tents. We know that if a population of genotypes can be expressed as a chain, then the population will always support an unrooted perfect phylogeny. If a g-lattice is a tent structure, then we know that the $2/3$ of the time the population supports an unrooted perfect phylogeny. This result lead to the concept of an upward branching tree.

2 Upward Branching Tree

The definition of an upward branching tree is a structure that has no tents and the g-lattice structure is represented by an upward facing tree. There are many benefits to this upward branching tree. For instance, if we know that the g-lattice is an upward branching tree, then we know several things about the population of genotypes. The upward branching tree can be broken up into chains and then the resulting haplotypes can then be easily determined. Another benefit to the upward branching tree is we know the upper bound on the problem is $m + 1$ for a population of $m$ genotypes. However, the solution to the Pure Parsimony Problem could be lower if multiple versions of a particular genotype exist in a single population. A third advantage to the resulting g-lattice if it is an upward branching tree is we know that the population supports an unrooted perfect phylogeny. One apparent disadvantage to the upward branching tree is if there are multiple versions of a genotype. This possibility makes it possible for the resulting g-lattice structure that is not unique for a given population of genotypes; that is, there can be multiple versions of the g-lattice. This also makes it harder to have an exact upper bound on the problem.

3 Future Work

Now that we know the upper bound that exists for an upward branching tree, a more rigorous proof needs to be completed. This will be done using the fundamentals of Real Analysis. Since we know the g-lattice structures that support an unrooted perfect phylogeny, we want to determine if these particular structures do in fact exist in a given population of genotypes. We want to devise an algorithm to check if upward branching trees do exist in real populations. Biologically this would make sense since the Perfect Phylogeny Problem increases the biological significance to the Pure Parsimony Problem.