Introduction

- RNA consists of the following nucleotides: adenine (A), guanine (G), cytosine (C), and uracil (U).
- In RNA folding C bonds with G and A bonds with U.
- In an idealized RNA folding, there are no unpaired nucleotides.
- In this research, I ask and answer questions about different RNA sequences.

Definitions

- Plane Tree (p.tree):
  - A rooted tree with the vertices of each child having an order.
- Plane Tree Move (Figure 1):
  - Combines two edges on a tree to form a new pair of edges.
- Idealized RNA Folding (IRNAF):
  - A type of bonding in RNA where every nucleotide bonds with one other nucleotide. Additionally, the arcs that are used to represent these bonds do not cross (example in figure 2).

Questions

1. How many RNA sequences fold into a given plane tree?
2. Given an RNA sequence and a tree, which tree moves are allowed?
3. How many ways can a given RNA sequence of length 2n fold?
4. Which RNA sequences have the most/fewest foldings?

Question 1

- Proposition:
  - A given plane tree, with n edges, can be represented by $4^n$ RNA sequences.
- Proof Sketch:
  - A single edge can be represented as one of four bonds. As you add more edges you multiply 4 by the previous number of sequences which is the same as $4^n$.
- Future Use:
  - This could be used for the establishment of a relationship between multiple RNA sequences that fold into the same tree.

Question 2

- Proposition:
  - A valid move on RNA p.trees involve either two siblings with the same type of bond or a parent and child, where the child is in an reversed order of the parent.
- Proof Sketch:
  - See Figure 1.
- Future Use:
  - By knowing which tree moves are allowed, I am able to easily calculate how many moves a certain RNA plane tree can make.

Question 3

- Conjecture and Intuition:
  - An RNA sequence of length 2n that is not in the forms: $(CG)^n$, $(GC)^n$, $(AU)^n$, and $(UA)^n$, have at most $\binom{n}{2}$ foldings.
  - The p.trees with the most foldings and n edges can fold into a maximum of $\binom{n}{2}$ trees. Since no other sequences can fold into every tree these sequences are limited certain trees.
- Future Use:
  - By knowing the different ways that any given sequence can fold, I limit the number of plane trees that sequence can fold into.

Idealized RNA Folding

In this research, I ask and answer questions about different RNA sequences. In an idealized RNA folding, there are no unpaired nucleotides. RNA consists of the following nucleotides: adenine (A), guanine (G), cytosine (C), and uracil (U). The p.trees with the most foldings and n edges can fold into a maximum of $\binom{n}{2}$ trees. Since no other sequences can fold into every tree, these sequences are limited to certain trees.

Future Questions

- How many 2n letter RNA sequences can fold?
- What are the other sequences of RNA that have only one way of folding?

References


Acknowledgments

Thank you, Dr. Heitsch, the graduate students, and postdocs for presenting me with this opportunity and providing me with much support. Additional thanks to Dr. Cooper and Dr. Wilson for recommending me for this program. Finally, a thank you to the NSF DMS grant #1344199 and Georgia Tech for making this program possible.