Notes on Notes

This document tries to bring together some of the observations made while developing the Random Mutations puzzle bot scripts. Most of these observations came from seeing how the various versions of the Random Mutations scripts attempted to solve puzzles and noting how and where they sometimes failed.The basic strategy was try something simple, see where it fails, and (repeatedly) adapt - both in terms of bot development and how the bot runs.

In Random Mutations v0.10 on puzzle id=683941 we that see short stacks can be problematic. In particular they are sensitive to the location of loop boosts and the turning direction of closing pairs. What was not discussed here, but was brought up by Lee Bickle (hoglahoo) at EternaCon in his segment on single state puzzle solving is that sometimes with *In Silico* puzzles, a short stack joining two loops may not form because of the difference in energy between the target and natural states that may require changes to elements of the “natural state” misfold loop to raise its energy so that the target shape will be preferred instead. (This point was discussed, but not resolved in v0.43 Possibly significant Differences.)

In Sample of new draw output, we see that a boost can cause a slip of the hairpin shape when a G boost is adjacent to a stem with all Gs on the same side. This could be generalized to other boost situations (e.g. G boosts adjacent to purine rich stem sides or UU boosts adjacent to pyrimidine rich stems). It is also noted that “It seems like a RNA stem will rather slide where possible and become more straight, than have a bent internal loop” (although extensive evidence is not provided) and that flipping a pair (to break a slippable run of purines opposite pyrimidines) can often be enough to fix the mismatch.

There are several discussions in Benchmark v0.46: RRF. First is a pointer to the fact that the bot starts by assuming standard patterns for the closing stem pairs in multiloops. This closing pattern is a simple heuristic and could be made more sensitive to *in silico* energy models of loop energies, but for now at least the simple pattern works for a wide variety of puzzles. It also points out that many player puzzles use special tricks or common patterns such as “zig zags” which a puzzle bot may need to be trained to recognise. It also points out that heuristics such as initializing with “Cs on the inside of lopsided loops” which works well for some situations can have secondary effects such as also initializing the loop boosts on the inside of lopsided loops; in this case it didn’t hurt, but someone trying to stump a bot might take advantage of default behaviors.

In Saccharomyces Cerevisiae Partial we note that the orientation of multi-loop closures can help to stabilize a section, in particular for 4-stem multi-loops, having same opposite same pairings can make the loop closures stronger and more forgiving of stack similarities. Also noted is that the closing pair orientation for large hairpin loops on short stacks has a given preference for the *in silico* energy model, favoring a heuristic for the default closing and boosting pattern.

Multiloop Closing Pair Energies starts to look at the energies involved in various orderings of closing pairs in multi-loops to see how some orderings work better than others. Estimating Unboosted Multi-loop Energies takes the matter further and lays out 4 general categories of multi-loop closures that may prove useful for a heuristic strategy. (Note: this is dependent on the Vienna energy model used by the Puzzle Maker at the time.) While a skilled Puzzle Master might use more detailed knowledge of multiloop energies and boosted multi-loops to design tricky puzzles, a basic knowledge of multi-loop closing energies may help to solve the vast majority of puzzles without requiring close inspection of the modelled energy of each multiloop in the puzzle.

In Pre-Fold Flip Strategy we discuss the idea of doing some pre-checks for potential misfolds before attempting an initial trial fold. The idea is that it may be possible to detect certain likely to fail patternings and try to “fix” them in advance without wasting a computationally expensive (for larger puzzles at least) run of the folding engine. This document (along with Flower Power and Common RNA Folding Failure Modes) takes a look at some of the ways that simple *symmetrical* puzzles can fail. It shows that some puzzles are concerned about the mismatching of identical (or compatible?) stems and that mismatches occur frequently with multi-loops having many similar length short stems[1]. It also discusses Eli’s View on Hoglahoo’s strategy of looking at the energy differences in mismatching regions. Common RNA Folding Failure Modes takes this further, looking at more cases where symmetrical puzzle fail, showing that identical twin stems off the same multiloop can often cause problems, at least when opposite; twin stems that are adjacent don’t necessarily mismatch. However with the Flower Power puzzle we see that sometimes similar stems on adjacent multiloops can also cause mismatches.

[1] There are several reasons. Short stems have fewer possible solutions, hence there is a greater chance of identical stems (which can mispair) in designs with many short stems. Multi-loops (especially symmetrical ones) can place similar sized features in close proximity where the potential for misfolding is greater. Also, a multi-loop that places a stem+hairpin opposite a similar stem+hairpin has the potential to misfold into a structure containing only stems and inner-loops, which is often a lower energy solution in the energy model. However other misfolds are also possible such as where stems in adjacent multiloops mismatch and reduce the multiloop count by one; in this example two high energy tri-loops are converted into a lower energy 3-3 inner-loop, further enabling the mismatch. Ultimately it is always the case that the alternative folding is lower in energy, but it is often the matching stems that help to enable the mismatch. However the source of the mismatch isn’t always obvious, so simply avoiding matching stems may not always be enough.