**Better Together: Combining Analytical and Annealing**

**Methods for FPGA Placement**

Q1. In this paper proposed, how to combine the analytical and annealing method for FPGA placement?

Ans: Use VPR to pre-clusters the complex block types (e.g. DSP, IO, memory) and leaves the common primitive types (e.g. LUTs, FFs) for analytical placer to place individually. And provide architecture specific information to allow benchmarking.

Then, legalize the placement solutions from analytical placer via VPR’s legalizer. Finally, refine the solution via VPR’s annealer.

Q2. Why the convergence time of VPR SA stage can be reduced and further optimized the QoR in the method proposed in this paper?

Ans:

Because the placement solutions generated by aug-elfPlace already have good QoR, using it as the starting point for SA leads to faster convergence compared to directly operating VPR SA. Additionally, the use of Quench or Anneal further refines the QoR since the feature of SA.

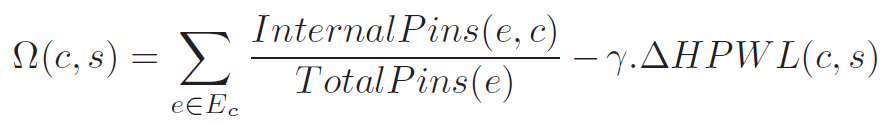
Q3. Propose possible reasons why adopting “partial macro representation” to handle carry chains can lead to smaller HPWL than using “net weighting” after legalization?

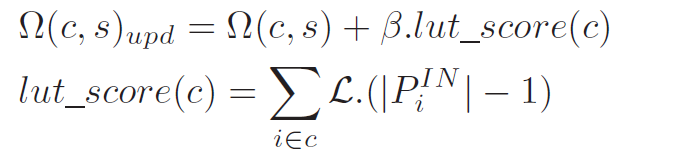
Ans:

Net weighting cannot guarantee the relative order of carry chain instances. During legalization, instances may be moved to satisfy constraints, increasing HPWL.

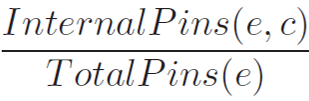
Partial Macro Representation treats the carry chain as a single macro, limiting movement and preserving order. This reduces displacement during legalization, minimizing the increase in HPWL.

Q4. Explain the goal of cluster candidate score and how to achieve with the following formula? In the definition of lut\_score(c)，L = 1 if instance i is LUT，L = 0 if instance i is FF. Why we only consider the number of input pins of LUT, can we also consider that of FF?





Ans: the goal of cluster candidate score is to evaluate whether a instance (such as LUTs or FFs) is suitable for packing into a specific cluster.

 this term encourage to enhance cluster internal connectivity ,  this term to optimize the wirelength,  and this term to consider the constraint of the Stratix -IV , i.e. max shared input of the LUTs/FFs.  
The reason why we only consider the number of input pins of LUT is that the amount of pins in a LUT is larger than that of a FF generally. Thus, we prioritize packing LUTs with more input pins together to reduce conflicts during the legalization process. If we also consider the amount of pins of FFs, it may result in FFs being prioritized for allocation into clusters instead of large LUTs that require more shared input resources.

Q5.

Implement the auction algorithm to assign 3 blocks to 4 sites and calculate the total cost.

Cost matrix between blocks and sites showing below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block \ Site | Site 1 | Site | Site 3 | Site 4 |
| Block 1 | 10 | 15 | 20 | 25 |
| Block 2 | 30 | 10 | 5 | 10 |
| Block 3 | 15 | 10 | 30 | 5 |

Ans: