

FlashTensor

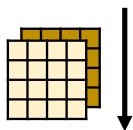
Tensor Property Identifier (§4)

Reduce Dependency: ...

Broadcast: ...

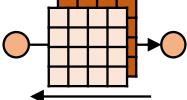
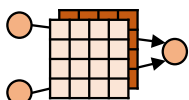
Size: ...

Value: ...



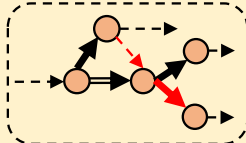
Aggregate

Forward

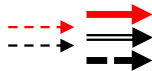


Backward

Tensor Program with Tensor Property

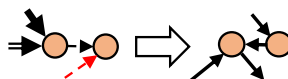


Different
Tensor Property

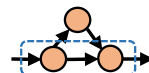


Tensor Property-Aware Optimization (§5)

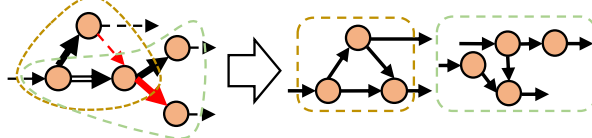
Algebraic Equivalent
Graph Transformation



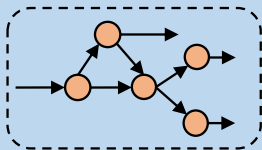
Non-Convex
Kernel Mapping



Lightweight Plan Searching



Input
Tensor Program



Optimized
Tensor Program

