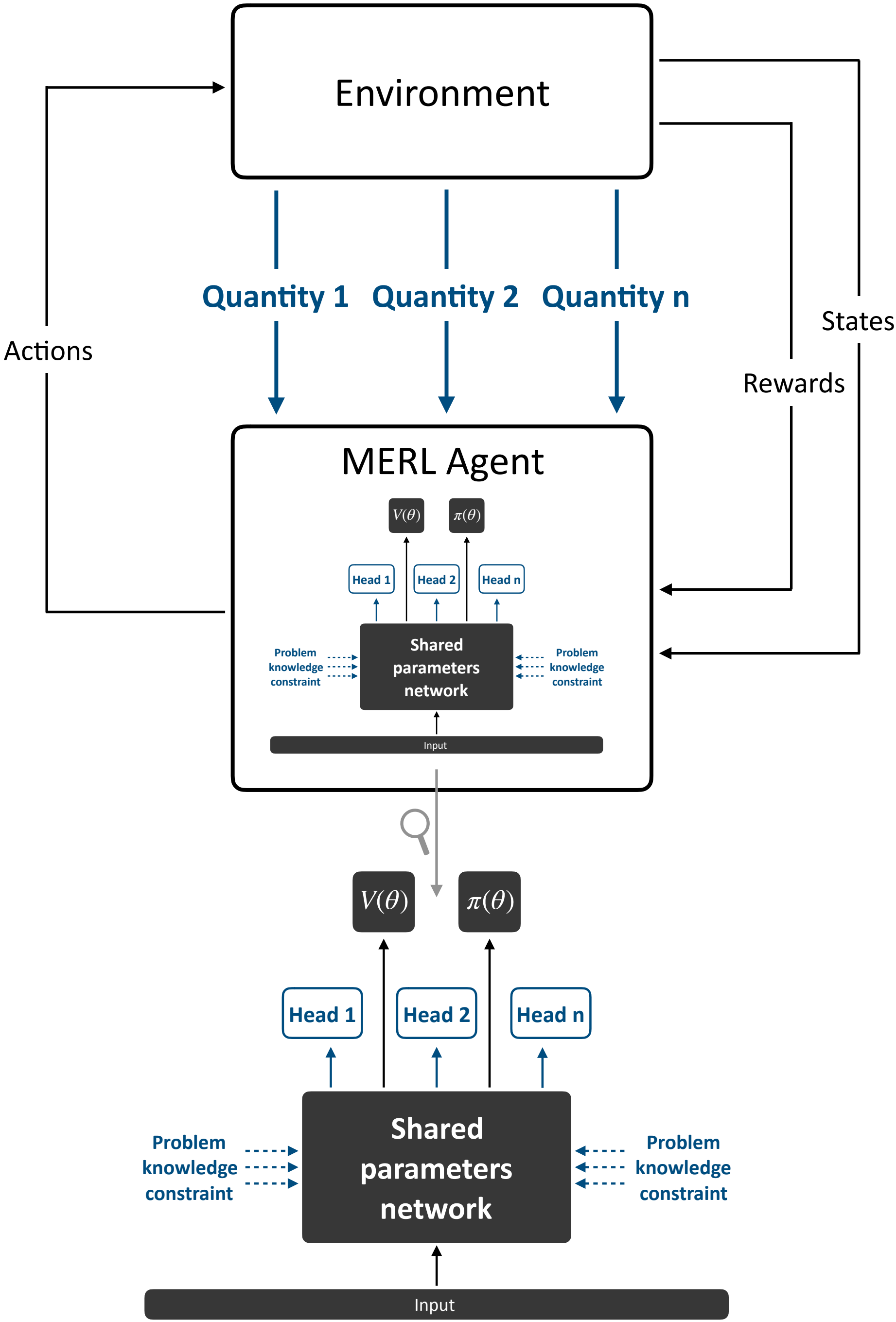
Environment
Quantity 1
Quantity 2
Quantity n
Actions
States
Rewards
MERL Agent
$V(\theta)$
$\pi(\theta)$
Head 1
Head 2
Head n
Shared parameters network
Problem knowledge constraint
Problem knowledge constraint
Input
$V(\theta)$
$\pi(\theta)$
Head 1
Head 2
Head n
Shared parameters network
Problem knowledge constraint
Problem knowledge constraint
Input