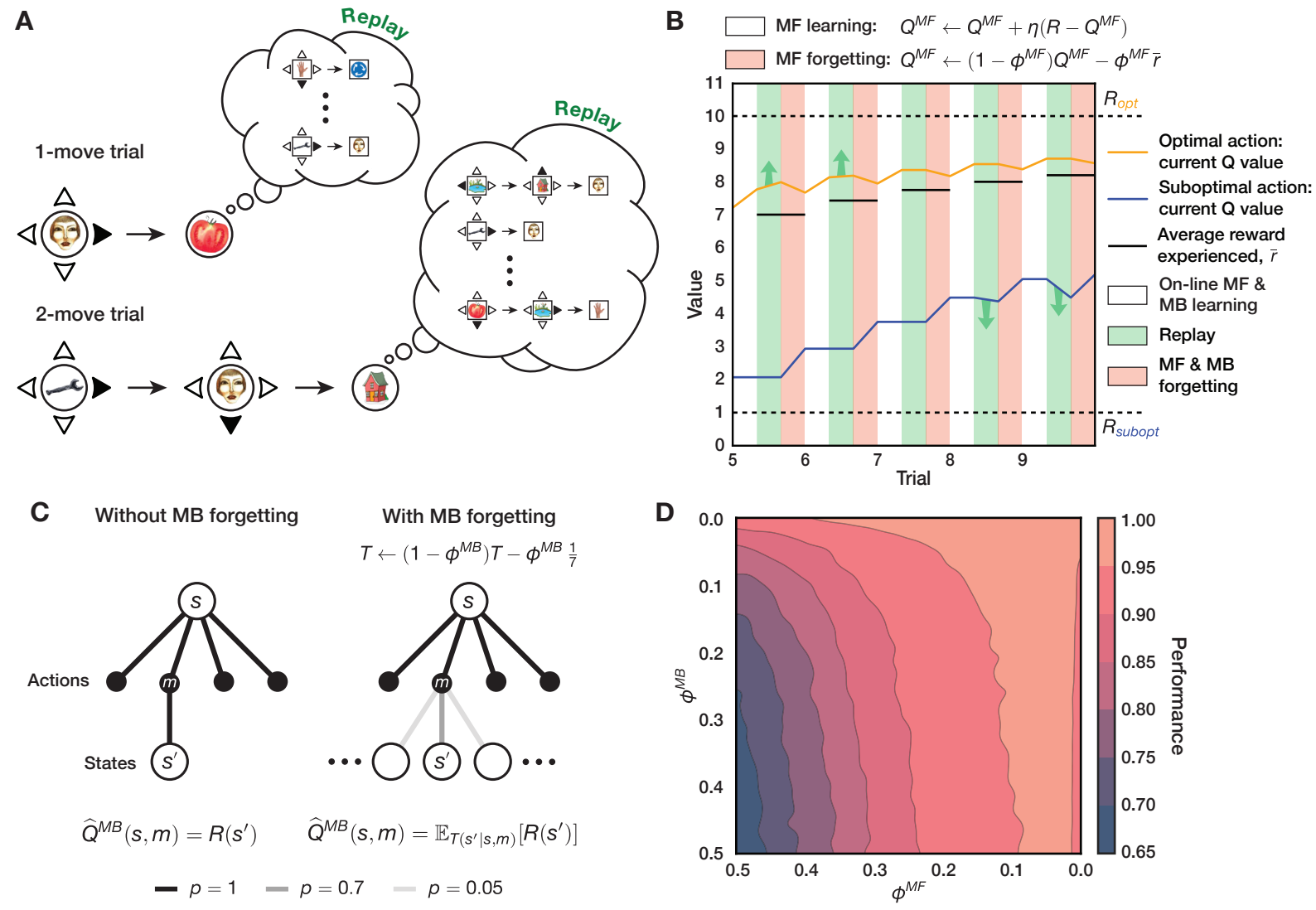
A
Replay
Replay
1-move trial
2-move trial
B
MF learning: $Q^{MF} \leftarrow Q^{MF} + \eta(R - Q^{MF})$
MF forgetting: $Q^{MF} \leftarrow (1 - \phi^{MF})Q^{MF} - \phi^{MF}\bar{r}$
Value
Trial
R_{opt}
R_{subopt}
Optimal action: current Q value
Suboptimal action: current Q value
Average reward experienced, $\bar{r}$
On-line MF & MB learning
Replay
MF & MB forgetting
C
Without MB forgetting
With MB forgetting
$T \leftarrow (1 - \phi^{MB})T - \phi^{MB}\frac{1}{7}$
Actions
States
$\widehat{Q}^{MB}(s,m) = R(s')$
$\widehat{Q}^{MB}(s,m) = \mathbb{E}_{T(s'|s,m)}[R(s')]$
$p = 1$
$p = 0.7$
$p = 0.05$
D
ϕ^{MB}
ϕ^{MF}
Performance