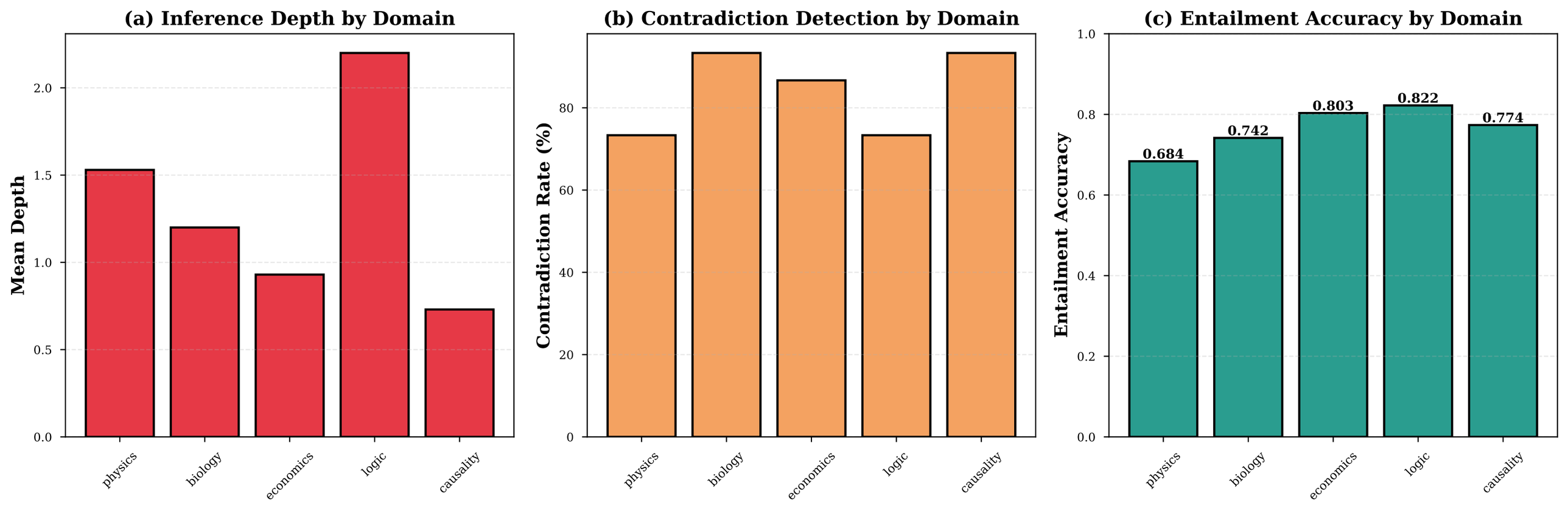
(a) Inference Depth by Domain
Mean Depth
0.0
0.5
1.0
1.5
2.0
physics
biology
economics
logic
causality
(b) Contradiction Detection by Domain
Contradiction Rate (%)
0
20
40
60
80
physics
biology
economics
logic
causality
(c) Entailment Accuracy by Domain
Entailment Accuracy
0.0
0.2
0.4
0.6
0.8
1.0
0.684
0.742
0.803
0.822
0.774
physics
biology
economics
logic
causality