

STRUCTURAL FRAGILITY

**Surface-Level
Semantic Reliance**



**Sharp Decision
Boundaries**



**Sensitivity to
Paraphrase / Wording**



SYSTEM FAILURE

**Output
Instability**



**Adversarial Trigger
Exploitation**