

Night Shift AI Studio —Optimized Full Talk Script (Bilingual CN/EN)

Grey

Usage Note

Scene “punchline”

Scene0 —Title / Claim (0:00–0:20)

EN: Hi everybody, good morning. My project is **Night Shift**.

EN: The key idea is simple: **Difficulty = computation budget**.

EN: With more budget, the engine sees further, becomes more stable, and avoids traps.

CN punchline : “ ”

EN (optional short pun): I call it *Night Shift* because it’s built with *knight*s — and I built it mostly at night.

Scene1 —Freeze: Same board, different choice (0:20–1:20)

EN: Let’s start with a quick question.

EN: Here is one position — **same board, same side to move**.

EN: The opponent’s queen is **hanging**.

EN: **Would you take the queen?**

(*Pause 2 seconds. Look at the audience.*)

EN: Now watch what two levels do with different budgets.

EN: Left: the low-budget level takes immediately — looks great.

EN: But a few moves later, it falls into a tactical trap.

EN: Right: the higher-budget level plays differently — it sees the trap first.

CN punchline : —

Scene2 —Difficulty Dial: four levels as budget presets (1:20–2:05)

EN: I implemented **four difficulty levels** by increasing computation budget and heuristic precision.

EN: Levels are not arbitrary names — they are **budget presets**.

- EN: **L1 Greedy (1-ply)**: very fast, very short-sighted.
- EN: **L2 Minimax + Alpha-Beta (depth 3)**: material-only eval + small random tie-break.
- EN: **L3 Practical engine**: alpha-beta + quiescence + transposition table, **0.6s/move**.
- EN: **Ultimate**: same as L3 but stronger eval (**PeSTO**) + **1.2s/move**.

CN punchline :

Scene3 —Knobs: why budget changes behavior (2:05–3:05)

EN: Difficulty is not one knob. It's four knobs:

1. EN: **Horizon** (depth / iterative deepening)
2. EN: **Efficiency** (alpha-beta, ordering, TT)
3. EN: **Eval richness** (material → positional → PeSTO)
4. EN: **Randomness** (controlled noise)

CN punchline : **Levels = presets of knobs**

Scene4 —Ladder: the behavior staircase (3:05–3:45)

EN: Think of the levels as a ladder:

- EN: L1 is fast and impulsive.
- EN: L2 sees simple tactics, misses deeper threats.
- EN: L3 is practical and stable.
- EN: Ultimate has stronger evaluation and more time, so it converts advantages more reliably.

CN punchline :

Scene5 —Search X-Ray: internal mechanism (3:45–5:05)

EN: What changes inside?

- EN: **PV** shows the best line.
- EN: **Alpha-beta** prunes irrelevant branches.
- EN: **Iterative deepening** stabilizes the best move under time limits.
- EN: **Quiescence** reduces horizon effect in tactical positions.

CN punchline : “ ”

Scene6—Evaluation Harness: fair, reproducible, scalable (5:05–6:05)

EN: To validate the ladder, I built an evaluation harness.

- EN: Same scoring: win 1 / draw 0.5 / loss 0.
- EN: Colors reversed for fairness.
- EN: Max 500 plies.

EN: Three protocols: M2M standard (100 games per pairing), H2M (10 participants), and time-scaled budgets.

CN punchline :

Scene7 —Evidence Wall + TSB strip (6:05–8:20)

EN: Evidence poster 1: round-robin scoreboard shows **separation**.

EN: Example highlights: **L2 vs L1 = 0.85, L3 vs L2 = 0.75**.

EN: Evidence poster 2: cost–strength curve shows **diminishing returns**.

EN: TSB strip shows the trend is stable across different budgets.

CN punchline :

Scene8 —Future Work (8:20–9:00)

EN: Future work: performance optimization, dynamic difficulty adjustment, and generalization to other board games.

CN punchline :

Scene9 —Closing / Handoff (9:00–9:30)

EN: To close: **Difficulty = computation budget**.

EN: I delivered a multi-level engine, a reproducible evaluation pipeline, and evidence of separation and diminishing returns.

EN: Thank you — now I can open the live demo. Q&A?

CN : /play