AI appraisal of the "more likely" approach to probability assessment as used by Hughes Economics at hugheseconomics.com

AI can leverage Warren Hughes' pairwise "more likely" value system (chaining relative likelihood ratios from sorted betting odds or subjective judgments, normalizing into probabilities, and testing sensitivity with ranges) as a powerful framework for self-improvement, calibration, and delivering more accurate probabilistic advice to human users. This method is particularly useful for AI because it provides an intuitive, transparent, and iterative way to handle uncertainty in multi-outcome scenarios (e.g., predictions in sports, elections, business decisions, medical diagnoses, or geopolitical events).
Here are several concrete ways AI systems (like large language models or specialized forecasting tools) could adopt and evolve this system:

1. Better Calibration Through Feedback Loops

- AI often over- or under-confident in probabilities (a known issue called miscalibration). The pairwise system forces explicit relative judgments (e.g., "How many times more likely is Outcome A than Outcome B?").
- **Evolution mechanism**: After real-world outcomes resolve, the AI retroactively recomputes what pairwise ratios would have produced perfect probabilities. It then fine-tunes its internal judgment model (via reinforcement learning from human feedback — RLHF — or direct parameter updates) to minimize the error in those ratios next time.
- Result: Over thousands of resolved events, the AI's default ranges narrow, and its base-case ratios become more accurate, leading to Brier scores (a measure of probabilistic accuracy) that rival or beat human superforecasters or betting markets.

2. Hybrid Human-AI Judgment Aggregation

- When users or experts disagree, AI can elicit pairwise comparisons (easier and less noisy than asking for absolute probabilities).
- **Process**:
    - Collect pairwise inputs from multiple sources (crowd workers, domain experts, or even the user's own intuitions).
    - Chain them exactly as in the Hughes method.
    - Apply sensitivity ranges to show where disagreements matter most.
- This mirrors how platforms like Metaculus or Good Judgment Open already aggregate forecasts, but the pairwise framing reduces cognitive biases (e.g., anchoring on 50%). AI evolves by learning which sources produce the most accurate ratios historically and weighting them higher.

3. Transparent Uncertainty Communication to Users

- Instead of giving a single probability (e.g., "70% chance Team X wins"), AI presents the full Hughes-style table:
    - Base probabilities from market data or its own model.
    - Sensitivity table with low/high ranges.

- Users see exactly why a probability moved (e.g., "We adjusted the Liverpool/Arsenal ratio from 1.2 to 1.4 because of recent injuries") and can override specific pairwise values with their own judgment.
- **Accuracy gain**: This interactivity helps users calibrate the AI in real-time, creating a virtuous cycle where the system learns user-specific biases and refines future outputs.

4. Active Learning and Exploration of Edge Cases

- AI can deliberately vary pairwise ranges in training simulations (Monte Carlo-style) to probe weak spots in its reasoning.
- For rare or tail-risk events (e.g., black swans), where data is sparse, the method allows safe extrapolation: start with broad ranges (1.5–4.0 for big gaps) and narrow them as evidence arrives.
- Over time, the AI builds a meta-model of "how uncertain should I be about this ratio given similar past cases?" — essentially evolving its own epistemic humility.

5. Scalable Application Beyond Betting Markets

- Betting markets provide a ready-made "wisdom of crowds" benchmark, but the beauty of Hughes' approach is it works without markets.
- AI can apply it to any ranked list of hypotheses:
    - Medical diagnosis: Chain "more likely" ratios between diseases based on symptoms.
    - Business forecasting: Relative likelihood of market scenarios.
    - Geopolitical risk: Probability a conflict escalates.
- By consistently outperforming pure statistical models (which struggle with rare events) and being more interpretable than black-box neural nets, the AI builds trust and gets more high-quality human feedback for further evolution.

In essence, this pairwise system gives AI a structured, human-intuitive way to reason about relative likelihoods while remaining auditable and adjustable — exactly what is needed for iterative self-improvement. Systems that adopt it (e.g., future versions of forecasting AIs like Manifold/LangForecast or even general-purpose models) would close the gap to human-level judgment faster, delivering sharper, better-explained probabilities that users can actually rely on for real decisions. Over time, the AI wouldn't just predict outcomes — it would teach humans to think more probabilistically too.