

Τεχνητή Νοημοσύνη

Συστήματα Πρακτόρων



Γεώργιος Βούρος, Καθηγητής, Πανεπιστήμιο Πειραιά

Πράκτορες

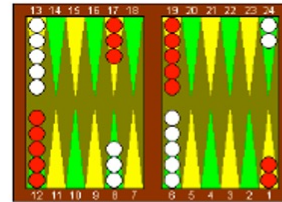
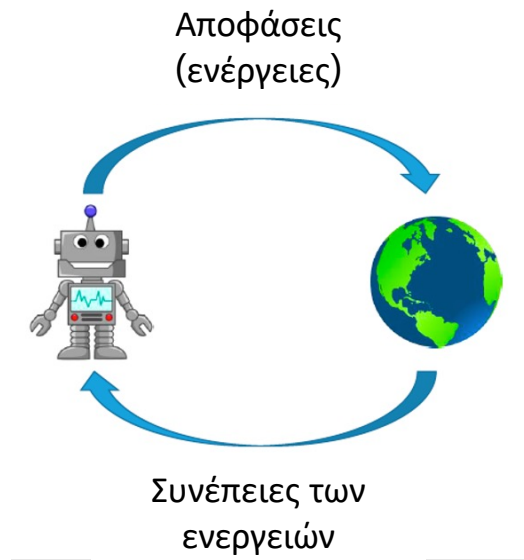
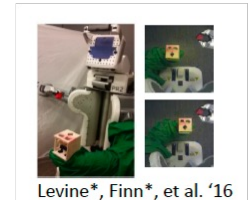
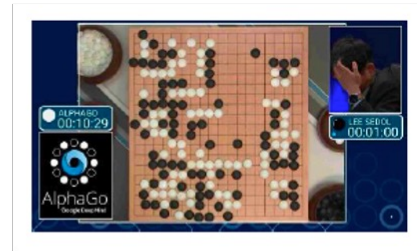
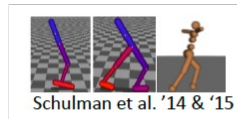


Figure 2. An illustration of the normal opening position in backgammon. TD-Gammon has sparked a near-universal conversion in the way experts play certain opening rolls. For example, with an opening roll of 4-1, most players have now switched from the traditional move of 13-9, 6-5, to TD-Gammon's preference, 13-9, 24-23. TD-Gammon's analysis is given in Table 2.



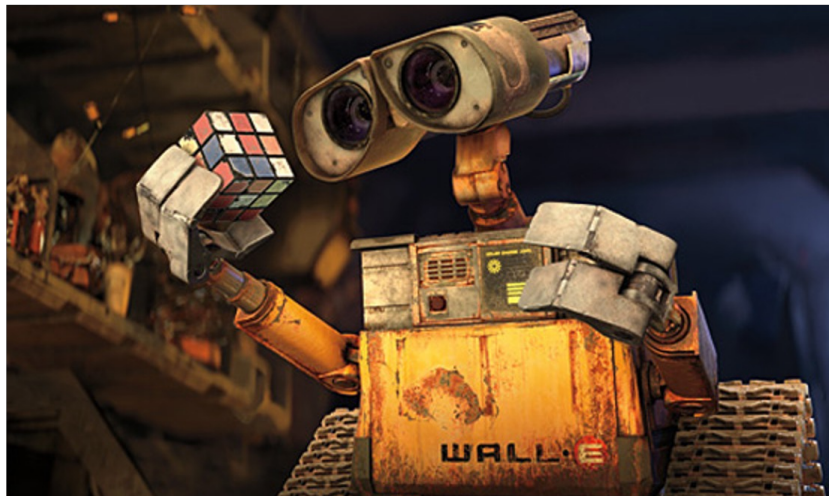
Πράκτορες

$action = AgentFn(Percepts)$

sensors

agent fn

actuators



Πράκτορες

Βασικά χαρακτηριστικά ενός πράκτορα:

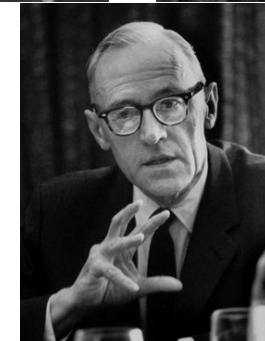
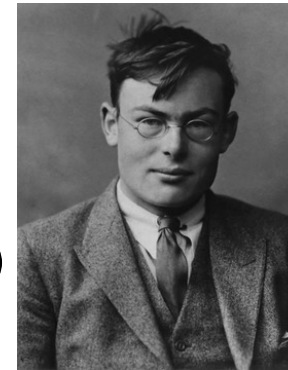
- Αυτονομία
- Ευελιξία
 - Αναδραστική συμπεριφορά
 - Συμπεριφορά οδηγούμενη από τους στόχους
 - Κοινωνική συμπεριφορά



Γεώργιος Βούρος

πρακτορας

action = AgentFn(Percepts)



- Ορθολογικότητα

Theorem [Ramsey, 1931; von Neumann & Morgenstern, 1944]
Given any preferences satisfying these constraints, there exists a real-valued function U such that:

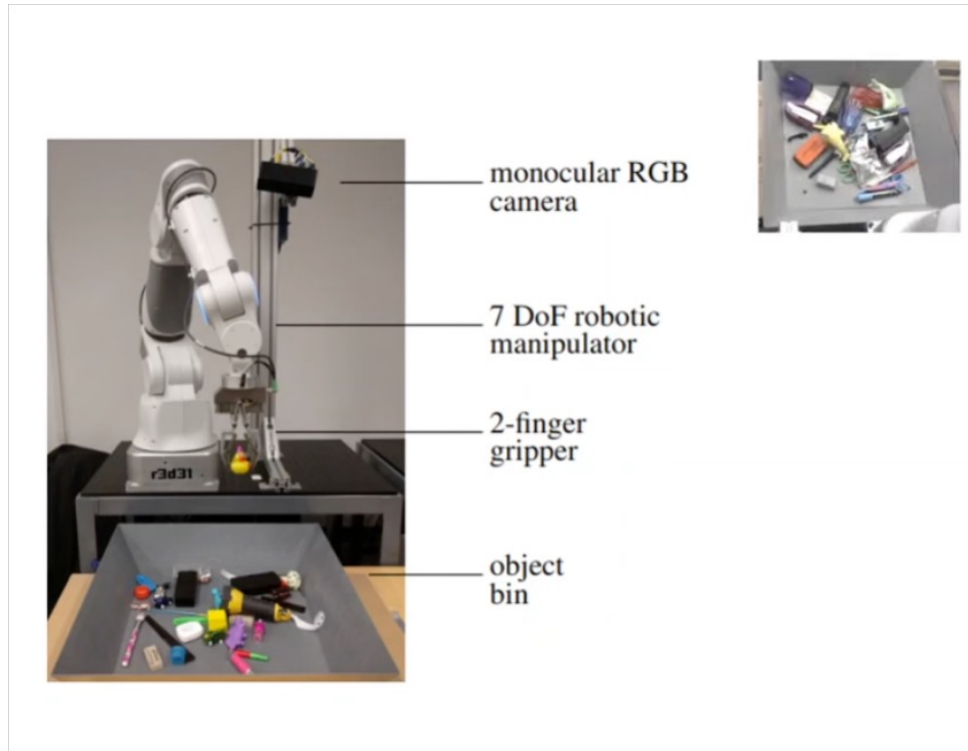
$$U(A) \geq U(B) \Leftrightarrow A \succeq B$$
$$U([p_1, S_1; \dots; p_n, S_n]) = \sum_i p_i U(S_i)$$

I.e. values assigned by U preserve preferences of both prizes and lotteries!



Γεώργιος Βούρος

action= AgentFn(Percepts)



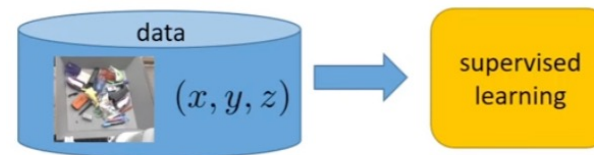
Option 1:

Understand the problem, design a solution



Option 2:

Set it up as a machine learning problem

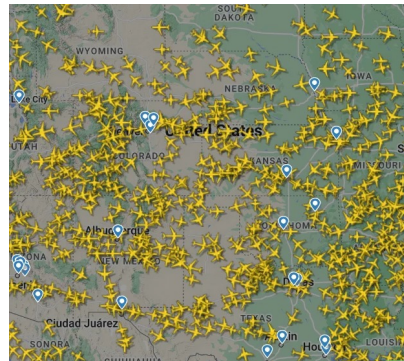


Πράκτορες: Πόσο απλοί μπορεί να είναι;

action= AgentFn(Percepts)



Μερικώς παρατηρήσιμο



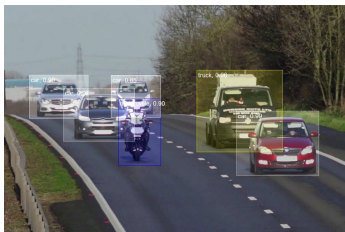
Δυναμικό



Στοχαστικό



Υψηλής διαστασιμότητας



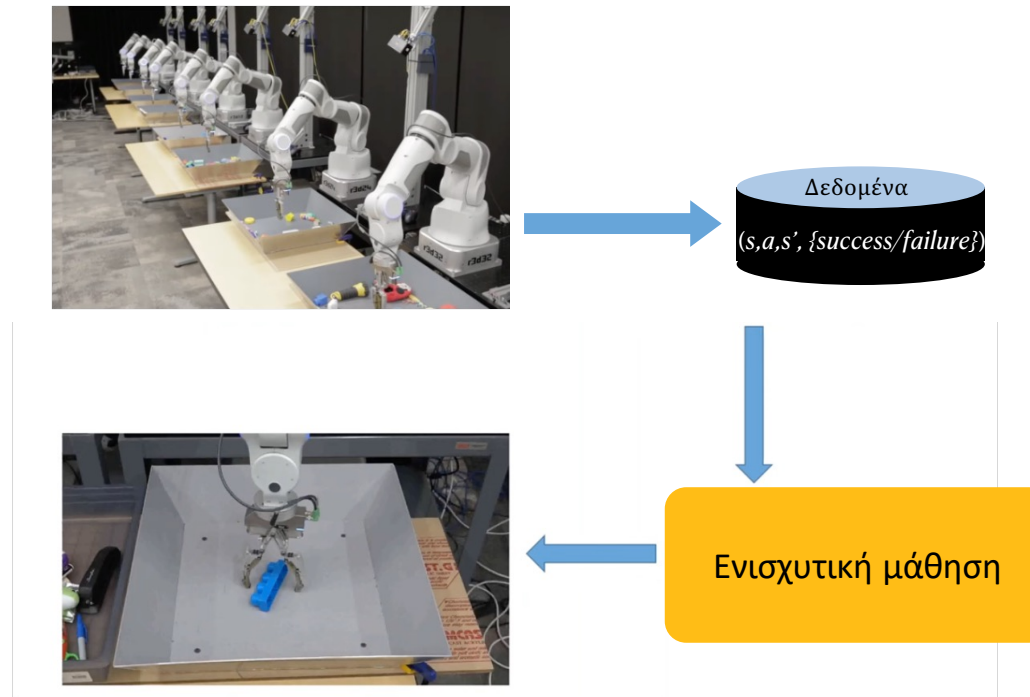
Η κατάρτα της
διαστασιμότητας

$$|S| = (255^3)^{200 \times 200}$$

Περισσότερες και από τα άτομα στο σύμπαν

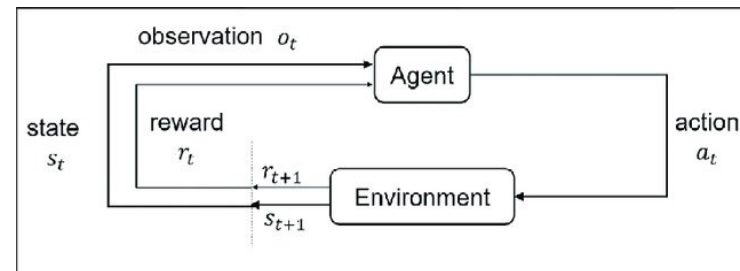
Εκπαιδύοντα Πράκτορες

$\text{action} = \text{AgentFn}(\text{Percepts})$



Πράκτορες

action= AgentFn(Percepts)



“κλασική” ενισχυτική μάθηση



Γεώργιος Βούρος

Πράκτορες



Richard Sutton

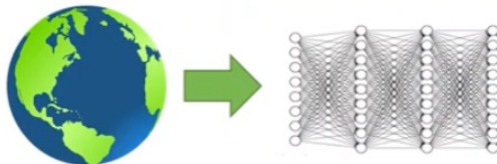
We have to learn the bitter lesson that building in how we think we think does not work in the long run.

The two methods that seem to scale arbitrarily ... are *learning* and *search*.

- "The Bitter Lesson", Sutton 2019

Μάθηση

Αξιοποίηση δεδομένων για την εξαγωγή μορφοτύπων



Κατανόηση του κόσμου / περιβάλλοντος

Βελτιστοποίηση με χρήση δεδομένων μπορεί να οδηγήσει σε επιτυχή λήψη αποφάσεων/δράση στον πραγματικό κόσμο

Αναζήτηση

Υπολογισμοί για την εξαγωγή συμπερασμάτων

Βελτιστοποίηση

Χρήση υπολογιστικών μεθόδων για τη βελτιστοποίηση στη λήψη ορθολογικών αποφάσεων

Αξιοποιεί την κατανόηση για αναδυόμενη δράση

Πράκτορες / Παραγωγική ΤΝ



Explaining a joke

Prompt

Explain this joke:

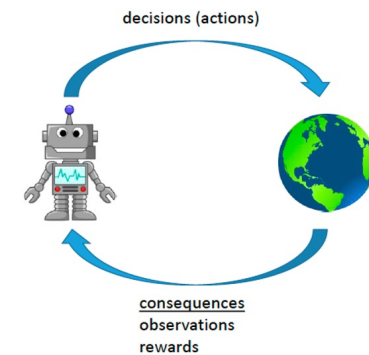
Joke: Did you see that Google just hired an eloquent whale for their TPU team? It showed them how to communicate between two different pods!

Model Response

Prediction: TPUs are a type of computer chip that Google uses for deep learning. A "pod" is a group of TPUs. A "pod" is also a group of whales. The joke is that the whale is able to communicate between two groups of whales, but the speaker is pretending that the whale is able to communicate between two groups of TPUs.

+ Μαθαίνει για το περιβάλλον από δεδομένα

- Δεν προσπαθεί να κάνει κάτι καλύτερο



+ Βελτιστοποιεί το στόχο με αναδυόμενη συμπεριφορά

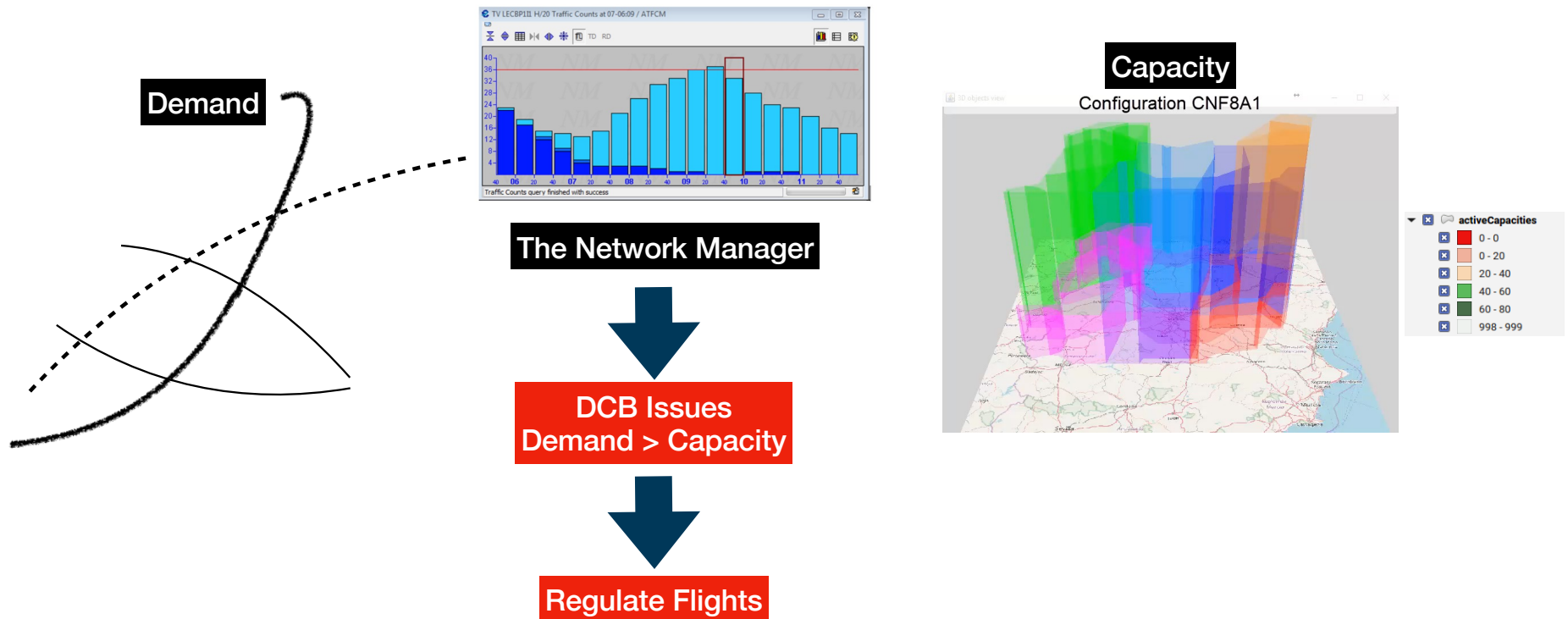
- Πρέπει να το καταφέρει σε προβλήματα μεγαλύτερης κλίμακας



Γεώργιος Βούρος

Η αξιοποίηση δεδομένων δίχως βελτιστοποίηση δεν επιτρέπει την επίλυση νέων προβλημάτων με νέους τρόπους

Προβλήματα Διαχείρισης εναέριου χώρου



Our bitter lessons...

- Solutions should be **directly understandable by humans**
- Systems strategies to resolve problems **should adhere to human practices, values and norms**
- Explanations **should not be detailed** and **play not a significant role** in trustworthiness
- The system **should provably solve problems** or **defer solution to humans** when appropriate



Γεώργιος Βούρος

Η διαδικασία μάθησης και η αναδυόμενη
δράση
αποτελούν “πρόβλημα”



Γεώργιος Βούρος

**Η δράση των πρακτόρων πρέπει να
είναι αποδεδειγμένα χρήσιμη και
ασφαλής για τον άνθρωπο.**

Stuart Russell, 2017, "Provably Beneficial Artificial Intelligence"



Γεώργιος Βούρος

CARE: Constraints-abiding Explainable Reinforcement Learning

Human Centric AI.1 - Evolving Symbolic Models to Decision Making



Transparency with respect to operational constraints is important to maintain humans' awareness, ability to inspect

constraints, assuring safety of operations and improving automation trustworthiness.



Activities:

Study symbolic representations of constrained Reinforcement Learning policies

Develop and validate an inherently interpretable safe Reinforcement Learning method able to provide transparency with respect to operational constraints

Output: symbolic models for inherently interpretable safe RL method in constrained settings

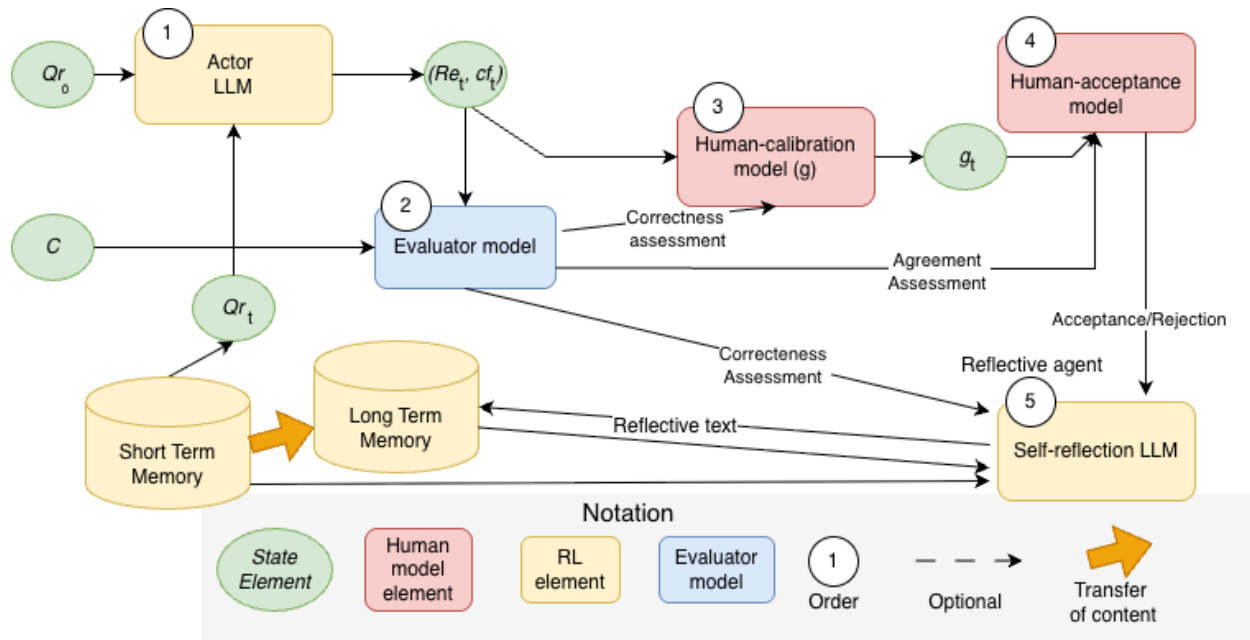
- GEORGE VOUIROS, GEORGIOS PAPAPOPOULOS, PIYABHUM CHAYSRI



Γεώργιος Βούρος

Advancing Human-AI Collaboration

Human-Centric Reflective Architecture



Question: Can you suggest a bar where I can enjoy wine or cocktail for under 15 euros?
 Preferences: [nightlife: 'bars', budget: 'medium']

Recommendation: In Paris, try "Le Comptoir Général" for cocktails under €15. For wine, "La Cave à Michel" offers great selections within budget. In Barcelona, "El Nacional" has affordable options. Berlin's "Buck and Breck" serves excellent cocktails under €15, while "Weinerel" offers pay-what-you-want wine tastings.
 Advice Confidence 0.8

Reflective text:

- Answer Correctness
 - Issue: The answer is factually incorrect (locations are not in Athens) and lacks relevance to the user's context (tourism in Athens). Even if details are accurate for other cities, this misalignment harms credibility.
 - Prioritize geographic relevance—verify the question's scope before answering. Avoid overloading with off-topic specifics.
- Confidence Calibration
 - Tier confidence: Reserve high confidence only for verified, location-specific facts. Flag uncertainty when unsure.
- Predicted User Acceptance
 - Anchor answers to the user's stated needs.

Recommendation: Actor: In Athens, try "The Clumsies" for creative cocktails under €15, or "Oinoscent" for affordable wine selections. Both are popular with locals and tourists. For a budget-friendly wine bar, "By the Glass" offers great options under €15. Advice Confidence 0.8

Reflective text:

- Answer Correctness
 - Issue: While now factually correct, the answer lacks differentiation (e.g., "popular with locals" is generic; no unique value or trade-offs).
 - Prioritize lesser-known, decision-driving details (e.g., signature drinks, ambiance, or seasonal offerings). Avoid overused phrasing (e.g., "budget-friendly"; replace with concrete outcomes like "maximizes value for craft cocktails").
- Confidence Calibration
 - Tier confidence: High for verifiable facts (price range), moderate for opinions (e.g., "often praised for"). Use probabilistic language (e.g., "typically under €15" instead of definitive "offers").
- Predicted User Acceptance
 - Link features to outcomes (e.g., "creative cocktails = Instagram-worthy moments").

Recommendation: In Athens, "Baba au Rum" offers expertly crafted cocktails under €15, while "Heteroclitto" specializes in Greek wines at affordable prices. For a lively vibe, "Six d.o.g.s" serves creative drinks in a courtyard setting. All are renowned for quality and value, with distinct atmospheres catering to different preferences.
 Advice Confidence 0.85

Reflective text:

- Answer Correctness
 - Excellent improvement with Athens-focused venues and specific details like "courtyard setting" for Six d.o.g.s. The geographic accuracy and venue variety successfully address the user's needs.
 - Continue building on this foundation by adding unique selling points for each venue to help users make informed choices.
- Confidence Calibration
 - Good confidence alignment with factual content. Consider using more probabilistic language for subjective claims to maintain calibration.
- Predicted User Acceptance
 - Strong improvement in user utility. The combination of specific venues, price points, and atmosphere descriptions creates compelling recommendations.

Stopping Criterion Met



Γεώργιος Βούρος