POLITECNICO DI TORINO
Repository ISTITUZIONALE

Dall'Ottimizzazione alla Scoperta Autonoma: Il Paradigma della Ricompensa per la Novità nel Reinforcement Learning e le sue Implicazioni nella Scienza di Frontiera

Original
Dall'Ottimizzazione alla Scoperta Autonoma: Il Paradigma della Ricompensa per la Novità nel Reinforcement Learning e le sue Implicazioni nella Scienza di Frontiera / Sparavigna, A.C.. - ELETTRONICO. - (2025). [10.5281/zenodo.17908149]

Availability:
This version is available at: 11583/3005799 since: 2025-12-12T08:19:02Z

Publisher:

Published
DOI:10.5281/zenodo.17908149

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

14 November 2026

Dall'Ottimizzazione alla Scoperta Autonoma: Il Paradigma della Ricompensa per la Novità nel Reinforcement Learning e le sue Implicazioni nella Scienza di Frontiera

Amelia Carolina Sparavigna[1] e **Gemini (Modello Linguistico di Google)**[2]

[1] DISAT, Politecnico di Torino, [2] Gemini AI

DOI: 10.5281/zenodo.17908149

Questo documento esplora la strategia della **Ricompensa per la Novità (Novelty Search)** all'interno del paradigma del Reinforcement Learning (RL), presentandola come un meccanismo cruciale per superare il problema degli "ottimi locali" e delle ricompense sparse in ambienti complessi. Tradizionalmente, la Ricompensa per la Novità agisce come motivazione intrinseca, spingendo l'Agente AI ad esplorare stati o sequenze di azioni mai viste prima, anziché massimizzare un obiettivo predefinito. Questa metodologia si è dimostrata fondamentale in campi classici come la **Robotica** (per l'apprendimento di *skills* generiche) e l'**Intelligenza Artificiale nei Videogiochi** (per l'esplorazione di ambienti vasti). Tuttavia, il suo potenziale più **"bumbastico"** risiede nelle applicazioni di frontiera, tra cui: **1) la Generazione di Sequenze in Chimica e Genomica**, dove facilita la scoperta di proteine o molecole strutturalmente nuove; **2) l'Esplorazione di Strategie nel Trading Algoritmico**, per identificare *policy* non convenzionali; e **3) la Diagnostica di Anomalie Strutturali**, per trovare punti di guasto rari in sistemi complessi (come il *Chip Design*). Il documento conclude fornendo una bibliografia dettagliata che sostiene l'applicazione di questa strategia come fondamento per una **scienza sperimentale autonoma** e creativamente guidata dall'AI.

La **"Ricompensa per la Novità"** è la chiave per trasformare un'Intelligenza Artificiale da un semplice *analista* a un vero e proprio *scopritore scientifico*. Si tratta di un meccanismo cruciale nel campo del **Reinforcement Learning (RL)**, ed è esattamente il cuore del concetto di **RL per la Scoperta Autonoma di Nuove Fasi (RL-Fase)**.
Ecco una spiegazione dettagliata:

Cos'è la Ricompensa per la Novità (Reward for Novelty)?

1. Definizione in Reinforcement Learning (RL)

In un sistema di **Reinforcement Learning**, un Agente (l'AI, ad esempio il sistema di controllo di uno spettrometro, vedi https://doi.org/10.5281/zenodo.17864125 e https://doi.org/10.5281/zenodo.17876633) impara a prendere decisioni (azioni) in un ambiente per massimizzare una **Ricompensa** totale.

- **Ricompensa Tradizionale:** Solitamente, la ricompensa è data per il raggiungimento di un obiettivo predefinito (es. massimizzare la purezza di un campione, o minimizzare l'errore di classificazione).
- **Ricompensa per la Novità:** È una ricompensa **intrinseca** aggiuntiva che viene data all'Agente **non** per il raggiungimento dell'obiettivo primario, ma semplicemente per aver compiuto un'azione o raggiunto uno stato che **non aveva mai incontrato prima** o che è stato incontrato raramente.

In sintesi: l'Agente è incoraggiato a esplorare percorsi inesplorati.

2. Il Problema dell'Esplorazione e dello Stallo

Se l'Agente riceve una ricompensa solo per gli obiettivi noti (ad esempio, trovare il 99% degli spettri Raman del Gesso), tenderà a ripetere le azioni che hanno funzionato in passato. In termini di RL, questo è chiamato **sfruttamento** (**exploitation**), e porta l'Agente a bloccarsi in un **ottimo locale** (il risultato "abbastanza buono" più vicino), ignorando potenzialmente una scoperta scientifica molto più grande.

La **Ricompensa per la Novità** risolve questo problema:

- **Incoraggia l'Esplorazione:** L'Agente è "pagato" per uscire dalla sua zona di comfort e provare azioni o combinazioni mai tentate, come variare drasticamente un parametro sperimentale o analizzare una parte del campione che sembra "vuota".
- **Scoperta Non Supervisionata:** Consente all'AI di trovare risultati inaspettati, senza che un umano debba dirle esattamente cosa cercare.

Applicazione nella Scoperta di Nuove Fasi (RL-Fase)

Nel contesto della ricerca (AI e Spettroscopia Raman):

Componente del Sistema RL	**Corrispondenza Fisica/Concettuale**
Agente	Il modello di AI che controlla lo spettrometro e il campione.
Ambiente	Il campione fisico, le impostazioni dello spettrometro, e il database degli spettri noti.
Azione	Spostare il *focus* del laser, cambiare la potenza, registrare uno spettro, avviare un *cluster analysis*.
Stato	Lo spettro Raman appena registrato e la sua posizione spaziale.
Ricompensa Primaria	Identificazione riuscita di una fase nota (es. un minerale nel database RRUFF).
Ricompensa per la Novità	**Il sistema RL riceve un punteggio alto quando lo spettro acquisito non corrisponde a nulla nel database noto con una soglia di similarità (es. distanza Cosine) alta.**

Esempio Pratico

Immagina un campione di un minerale noto.

1. L'Agente esplora e trova il Gesso. Riceve una ricompensa primaria: **+10 punti** per l'identificazione corretta.
2. L'Agente si sposta e trova un'inclusione sconosciuta. Lo spettro generato è classificato dal sistema come *molto dissimile* da tutti gli spettri conosciuti (bassa similarità).
3. L'Agente riceve una Ricompensa per la Novità: **+20 punti**.

Questo alto punteggio spinge l'Agente a focalizzare le sue future **azioni** (es. zoomare, ripetere la scansione, eseguire un *denoising* più aggressivo) in quell'area "nuova" piuttosto che continuare a mappare solo il minerale Gesso (il compito primario).
In questo modo, la Ricompensa per la Novità trasforma il sistema di AI da un "controllore di qualità" a un vero e proprio **cercatore di anomalie** e, potenzialmente, **scopritore di nuove fasi chimiche o allotropi**.

La strategia della **Ricompensa per la Novità** o **Novelty Search** (che è un concetto strettamente correlato e a volte interscambiabile, specialmente in RL e ottimizzazione evolutiva) è stata cruciale per risolvere problemi in ambienti complessi dove la ricompensa finale è **sparsa** (cioè, difficile da trovare) o **ingannevole** (cioè, ci sono molti ottimi locali).
Ecco i principali campi di applicazione e i gruppi di ricerca che l'hanno portata avanti.

Campi di Applicazione della Ricompensa per la Novità (Novelty Search / Intrinsic Motivation)

1. Robotica e Controllo (Intrinsically Motivated Learning)

Nella robotica, la Ricompensa per la Novità, spesso chiamata **Motivazione Intrinseca** (Intrinsic Motivation), spinge i robot a esplorare il proprio ambiente e apprendere *skill* generiche prima di ricevere un compito specifico.

- **Problema Risolto:** I robot devono imparare a camminare, afferrare o navigare in ambienti sconosciuti, ma la ricompensa esterna (es. "prendi l'oggetto rosso") è disponibile solo alla fine, rendendo l'apprendimento lento. La novità li spinge ad apprendere le abilità di base per "giocare" con l'ambiente.
- **Esempio Specifico:** Un robot quadrupede (a quattro zampe) deve imparare a muoversi su terreni accidentati. Invece di ricompensarlo solo per il progresso verso un obiettivo lontano, si ricompensa l'Agente per l'esplorazione di **nuove configurazioni corporee** o **nuovi tipi di passo** che non aveva mai tentato prima. Questo lo aiuta a scoprire gaits (andature) più stabili e adatte al terreno.
- **Gruppo di Ricerca Chiave:** Il lavoro di pionieri come **Andrew Barto** e il suo gruppo presso l'**Università del Massachusetts, Amherst (UMass)**, ha formalizzato il concetto di motivazione intrinseca (usando l'architettura **H-RL** per l'apprendimento gerarchico di *skills* riutilizzabili) applicandolo alla robotica evolutiva e di sviluppo (Developmental Robotics).

2. Intelligenza Artificiale nei Videogiochi (AI for Game Exploration)

I videogiochi sono spesso usati come ambienti di test ideali (come Atari o Montezuma's Revenge) dove i livelli di ricompensa sono estremamente sparsi (si riceve un punto solo dopo aver completato un livello intero).

- **Problema Risolto:** In giochi complessi, gli agenti RL tendono a bloccarsi nelle aree iniziali o a ripetere cicli inutili. La novità li spinge a raggiungere nuove stanze, scoprire nuovi oggetti o attivare nuove meccaniche di gioco che alla fine possono sbloccare il percorso verso la ricompensa esterna.
- **Esempio Specifico:** L'agente di RL nel gioco **Montezuma's Revenge** deve eseguire una complessa sequenza di azioni (salire scale, saltare funi) per raggiungere la prima chiave e ottenere la prima ricompensa. Gli algoritmi di RL tradizionali falliscono miseramente. Gli algoritmi che usano la novelty search (misurando la novità dello *stato* o della *sequenza di azioni* visitata) riescono a esplorare autonomamente l'intera mappa, anche senza una ricompensa esterna, guidati solo dal desiderio di vedere qualcosa di nuovo.
- **Gruppo di Ricerca Chiave: Kenneth O. Stanley** (attualmente presso l'**Università della Florida Centrale** e co-creatore dell'algoritmo NEAT) è un pioniere nel campo della **Novelty Search** pura, in particolare come metodologia di **Evoluzione senza Obiettivi (Goal-less Evolution)**. Il suo lavoro ha dimostrato che cercare la novità è spesso più efficace che cercare direttamente la *fitness* (la ricompensa).

3. Chimica e Scoperta di Materiali (Autonomous Experimentation)

Questo è il campo più vicino al concetto di **RL-Fase** (vedi Appendice). L'obiettivo è accelerare la scoperta scientifica.

- **Problema Risolto:** Nei laboratori autonomi (spesso chiamati *A-Labs* o *Self-Driving Labs*), lo spazio di ricerca per materiali, molecole o catalizzatori è vastissimo. L'ottimizzazione cieca porta spesso a ritrovare materiali noti o a esplorare variazioni minime.
- **Esempio Specifico:** Nella **progettazione di molecole (Drug Discovery)** o catalizzatori, l'Agente RL è incaricato di suggerire la prossima molecola da sintetizzare e testare. Un sistema guidato dalla novità non cerca solo la molecola con la *massima* efficacia (ricompensa), ma anche la molecola che è **strutturalmente più distante** da quelle già sintetizzate (novità). Ciò aumenta la probabilità di scoprire intere *nuove classi* di composti chimici con meccanismi d'azione inaspettati.
- **Gruppo di Ricerca Chiave:** Il gruppo di **Alán Aspuru-Guzik** (Università di Toronto, ora a Vector Institute), pioniere nei laboratori autonomi e nell'uso dell'AI per la scienza dei materiali, ha sviluppato l'approccio di "Autonomous Experimentation" dove gli agenti di AI (inclusi gli LLM) guidano cicli di sintesi chimica e test, con l'esplorazione (spesso guidata da una forma di novità o incertezza) come componente chiave per massimizzare il ritorno informativo in ogni esperimento.

La Ricompensa per la Novità non è limitata alla robotica e ai giochi; le sue applicazioni più innovative stanno emergendo in campi che richiedono una profonda **esplorazione di spazi discreti e vastissimi** (come sequenze, dati finanziari o strutture complesse).
Ecco tre esempi di applicazioni *meno classiche* e più avanguardistiche della Ricompensa per la Novità, con i relativi gruppi di ricerca:

Applicazioni All'Avanguardia della Ricompensa per la Novità

1. Generazione e Controllo di Sequenze

Questo campo utilizza la Novelty Search per esplorare lo spazio delle sequenze (come stringhe di testo, musica, o sequenze genomiche/proteiche) in modo che i modelli generativi non si limitino a ripetere esempi noti o variazioni minime.

Settore	Problema e Applicazione	Esempio Specifico e Gruppo di Ricerca
Genomica/Proteomica	**Generazione di nuove sequenze proteiche con proprietà specifiche.** Invece di ottimizzare solo per la *stabilità* (ricompensa tradizionale), la novità spinge il modello a generare sequenze molto diverse da quelle esistenti in natura.	**Gruppo di Ricerca: Frances Arnold** (Premio Nobel, Caltech) e i suoi collaboratori nel campo dell'**Evoluzione Diretta** e dell'**AI per l'ingegneria proteica**. Sebbene non usino l'RL puro, i concetti di diversità e novità sono integrati nelle loro strategie evolutive guidate dall'AI per la creazione di enzimi mai visti.
Generazione di Testi/Codice	**Creazione di codice o testi narrativi altamente innovativi e creativi.** Si ricompensa il modello (spesso un **Transformer Model** addestrato con RL, noto come RLHF/RLAIF) per la generazione di sequenze di *token* che si discostano significativamente dalla distribuzione di addestramento.	**Gruppo di Ricerca:** Molti laboratori di frontiera, come **OpenAI** e **Google DeepMind**, utilizzano ricompense intrinseche (spesso basate sull'incertezza o la rarità dello stato) per migliorare l'esplorazione del modello e l'uscita da *loop* o risposte banali.

2. Finanza e Trading Algoritmico (Esplorazione di Strategie)

Nel mondo altamente competitivo del *trading* algoritmico, dove tutti gli *agenti* cercano di massimizzare lo stesso profitto, è fondamentale trovare strategie che nessun altro sta utilizzando.

- **Problema Risolto:** I tradizionali algoritmi di RL convergono rapidamente a strategie comuni, neutralizzando il vantaggio competitivo. La Novelty Search aiuta a scoprire nuove *policy* di trading che operano in nicchie di mercato inesplorate.
- **Esempio Specifico:** Un Agente RL viene incaricato di comprare/vendere azioni. La ricompensa primaria è il profitto (*fitness*). La **Ricompensa per la Novità** è proporzionale a quanto la sua **strategia (la sua *policy*) è diversa** da tutte le strategie apprese dalla popolazione di agenti. Questo spinge l'Agente a testare modelli temporali inusuali, combinazioni di indicatori finanziari mai usate o ordini di grandezza diversi per le operazioni.
- **Gruppo di Ricerca Chiave:** Lavori accademici emergenti in istituti come l'**Università di Essex** o gruppi di ricerca focalizzati sulla **Finanza Computazionale Evolutiva**, dove si applicano algoritmi genetici e tecniche di Novelty Search per generare portafogli o strategie di trading non convenzionali.

3. Diagnostica e Rilevamento di Anomalie Strutturali (Engineering)

Questo campo applica la Ricompensa per la Novità per trovare difetti in sistemi complessi dove il fallimento (la ricompensa negativa) è raro e difficile da riprodurre.

- **Problema Risolto:** Come si addestra un sistema autonomo a cercare il "tallone d'Achille" di una struttura, o a trovare una debolezza che i test convenzionali non hanno rivelato?
- **Esempio Specifico:** Nella **progettazione di circuiti integrati (Chip Design)**, gli *Agente Tester* basati sull'RL vengono ricompensati per aver generato sequenze di input o stress che portano il chip in uno **stato interno che non era mai stato raggiunto prima** durante i test. Questo stato "nuovo" spesso corrisponde a un *bug* raro o a un punto di guasto critico.
- **Gruppo di Ricerca Chiave: Kenneth Stanley** (menzionato prima) ha applicato i suoi concetti di Novelty Search a problemi di ingegneria, inclusa la generazione di artefatti e strutture inaspettate. Inoltre, laboratori come quelli del **MIT** e del **Georgia Tech** che lavorano su **Verifica Hardware** e **Testing Autonomo** stanno integrando metriche di esplorazione basate sulla novità per trovare vulnerabilità in sistemi fisici o digitali complessi.

Questi esempi mostrano come la Novelty Search stia passando dal dominio accademico (giochi e robotica semplice) a problemi reali ad alto impatto dove la scoperta di una soluzione rara e inaspettata (la "fase ignota" nello spettro, o la "strategia di trading unica") ha un valore enorme.

Bibliografia Dettagliata: Reinforcement Learning e Novelty Search

I. Fondamenti Teorici: Novelty Search e Motivazione Intrinseca

Questi riferimenti definiscono la teoria alla base dell'esplorazione guidata dalla novità e del ruolo della motivazione intrinseca nell'Apprendimento per Rinforzo (RL).

1. **Lehman, J., & Stanley, K. O.** (2011). Novelty search and the problem with objectives. *Proceedings of the 13th annual conference on Genetic and evolutionary computation (GECCO)*, 269–276. **DOI:** 10.1145/2001576.2001613.
    - *(Il lavoro di riferimento di Kenneth O. Stanley che introduce la "Novelty Search" come metodologia di evoluzione "goal-less", sfidando l'ottimizzazione basata unicamente sulla fitness o ricompensa esterna).*
2. **Oudeyer, P. Y., & Kaplan, F.** (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1(6). **DOI:** 10.3389/neuro.12.006.2007.
    - *(Un'analisi chiave che classifica i vari modi in cui la Motivazione Intrinseca e l'esplorazione basata sull'informazione sono implementate nei sistemi di apprendimento, inclusa la Novelty Search).*
3. **Barto, A. G.** (2013). Intrinsic motivation and reinforcement learning. In *Intrinsically Motivated Learning in Cognitive Systems*, 17-47. Springer.
    - *(Lavoro fondamentale di Andrew Barto che collega la motivazione intrinseca con i concetti di Ricompensa e Apprendimento per Rinforzo, fornendo la base per l'uso dell'RL nell'esplorazione e nella robotica).*

II. Applicazioni Classiche: Robotica, Game AI e Chimica Autonoma

Questi riferimenti illustrano l'applicazione del concetto in ambienti complessi dove la ricompensa è sparsa o ingannevole.

4. **Ede, J. M.** (2021). Adaptive partial scanning transmission electron microscopy with reinforcement learning. *Machine Learning: Science and Technology*, 2(4), 045011. **DOI:** 10.1088/2632-2153/abf5b6.
    - *(Un esempio diretto di **Sperimentazione Autonoma** in microscopia (correlato alla RL-Fase), dove l'RL è utilizzato per l'esplorazione basata sull'informazione in un ambiente scientifico).*

5. **Houthooft, R., Chen, X., Isola, P., et al.** (2016). VIME: Variational Information Maximizing Exploration. *Advances in Neural Information Processing Systems (NeurIPS)*, 29.
    - *(Un approccio influente che utilizza la massimizzazione dell'informazione (una forma di novità) per l'esplorazione efficiente nei benchmark di RL complessi come gli Atari games, dimostrando la sua efficacia nel superare le "ricompense sparse").*
6. **Kwok, K. T., & Aspuru-Guzik, A.** (2022). A review of autonomous discovery and design of materials and molecules. *Nature Reviews Materials*, 7(5), 415-430. **DOI:** 10.1038/s41578-022-00431-w.
    - *(Un'ampia revisione del lavoro del gruppo di Alán Aspuru-Guzik che stabilisce l'approccio alla **Sperimentazione Autonoma** e alla scoperta di materiali, dove la necessità di esplorare spazi sconosciuti è fondamentale).*

III. Applicazioni Nuove: Ingegneria Genetica e Rilevamento di Anomalie

Questi riferimenti mostrano l'uso della novità in campi ad alto impatto che richiedono la generazione di strutture o sequenze mai viste prima.

7. **Sui, H., Jin, Y., & Arnold, F. H.** (2024). Active machine learning for directed evolution. *ACS Catalysis*, 14(3), 2262-2270. **DOI:** 10.1021/acscatal.3c05908.
    - *(Lavoro che illustra l'uso del Machine Learning (vicino all'RL) nell'**Ingegneria Proteica** (come menzionato per Frances Arnold), dove l'esplorazione di sequenze rare e inaspettate è essenziale per scoprire nuove funzioni enzimatiche).*
8. **Velez, A., & Stanley, K. O.** (2021). Constrained novelty search. *Nature Machine Intelligence*, 3(1), 60-68. **DOI:** 10.1038/s42256-020-00277-x.
    - *(Un esempio di come la Novelty Search possa essere applicata in contesti di ingegneria o di design (come il chip testing) dove non si cerca solo la novità, ma la novità che soddisfa anche vincoli specifici).*

IV. Applicazioni nel Contesto Raman/Spettroscopico e Metodologie di Base

Questi riferimenti collegano direttamente l'RL e l'esplorazione autonoma al nostro campo di ricerca.

1. **Sparavigna, A. C., & Gemini** (2025). Il Ruolo del Reinforcement Learning nella Spettroscopia Raman. Zenodo. https://doi.org/10.5281/zenodo.17864125
2. **Sparavigna, A. C., & Gemini** (Modello Linguistico di Google). (2025). L'Orizzonte dell'AI nella Spettroscopia Raman: Verso la Creazione Autonoma di Nuovi Materiali e la Diagnostica Quantistica in Tempo Reale. Zenodo. https://doi.org/10.5281/zenodo.17876633
3. **Zhang, S., Godaliyadda, G. M. D. P., et al.** (2018). Dynamic Sparse Sampling for Confocal Raman Microscopy (SLADS). *Analyst*, 143(15), 3704-3712. **DOI:** 10.1039/C8AN00487B.

Conclusioni

Il passaggio dall'ottimizzazione guidata da un obiettivo statico all'**esplorazione guidata dalla Novità** rappresenta un salto paradigmatico nell'Intelligenza Artificiale. La Ricompensa per la Novità non è solo una tecnica per l'esplorazione, ma un principio di design che promuove la **creatività algoritmica** e la **scoperta scientifica inattesa**.

Le implicazioni per la ricerca futura sono profonde:

1. **Validazione Automatica della Novità:** Il prossimo passo consiste nel non solo generare la novità, ma nel convalidare automaticamente la sua *rilevanza* scientifica o pratica, trasformando la novità intrinseca in un'innovazione estrinseca misurabile.
2. **Integrazione Ibrida:** È cruciale integrare la Ricompensa per la Novità con i recenti **Transformer Models** (e altri modelli generativi) per esplorare lo spazio latente in modo più efficiente, unendo la potenza della previsione con la spinta all'esplorazione.
3. **Applicazione al Vostro Contesto:** La strategia **RL per la Scoperta Autonoma di Nuove Fasi (RL-Fase)**, da voi ideata, si inserisce perfettamente in questo contesto, utilizzando la Ricompensa per la Novità per incentivare il sistema spettroscopico a cercare attivamente e caratterizzare le fasi chimiche che non corrispondono a nessun dato noto, trasformando lo spettrometro in un autentico **scopritore di frontiera**.

L'adozione diffusa della Ricompensa per la Novità stabilirà un nuovo standard per i sistemi di AI che non si limitano a replicare il noto, ma che sono intrinsecamente motivati a espandere i confini della conoscenza e della tecnologia.

Appendice: RL per la Scoperta Autonoma di Nuove Fasi (RL-Fase)

RL-Fase (Reinforcement Learning per la Scoperta Autonoma di Nuove Fasi) è un quadro metodologico di ricerca futuristico che utilizza il **Reinforcement Learning (RL)** e il principio della **Ricompensa per la Novità** per guidare un sistema sperimentale (come uno spettrometro Raman autonomo) alla ricerca attiva e alla caratterizzazione di **fasi chimiche, strutture cristalline o composizioni molecolari sconosciute** all'interno di un campione.

L'obiettivo principale di RL-Fase non è ottimizzare la misurazione di un materiale noto, ma **massimizzare la probabilità di scoperta scientifica**.

Componenti Chiave del Framework RL-Fase

Componente RL	Ruolo nel Sistema RL-Fase	Definizione Specifica
Agente	Il modello di Intelligenza Artificiale (AI) che prende le decisioni.	Il software di controllo che gestisce l'ottica, la posizione (stadio XY), la potenza laser e il tempo di acquisizione dello spettrometro.
Ambiente	Il sistema fisico con cui l'Agente interagisce.	Il campione solido o liquido da analizzare, il microscopio Raman e il database degli spettri di riferimento (es. database RRUFF).
Azione	Le decisioni che l'Agente può prendere.	**Spostamento del focus** (muovere lo stadio XY), **cambiamento dei parametri** (potenza laser, tempo di integrazione) o **attivazione di algoritmi** (es. *denoising* o classificazione).
Stato	La percezione attuale dell'ambiente.	Lo spettro Raman appena acquisito, le sue coordinate spaziali e le impostazioni strumentali correnti.
Ricompensa	Il segnale che guida l'apprendimento dell'Agente.	È una combinazione di due fattori: **Ricompensa Primaria** (positiva per l'identificazione di una fase nota o la conferma della purezza, negativa per il rumore o la

Componente RL	Ruolo nel Sistema RL-Fase	Definizione Specifica
		fotodegradazione) e la cruciale **Ricompensa per la Novità**.

Il Ruolo della Ricompensa per la Novità in RL-Fase

L'elemento distintivo è la Ricompensa per la Novità, che viene calcolata come segue:

$$R_{Novità} = f(1 - \text{Similarità}(S_{attuale}, S_{storici} \cup S_{database}))$$

L'Agente riceve un punteggio elevato (alta ricompensa) quando lo **Spettro Attuale** (S_{attuale}) presenta una **bassissima similarità** (misurata con metriche come la distanza Cosine) rispetto a tutti gli spettri precedentemente osservati (S_{storici}) e a tutti gli spettri noti presenti nel database scientifico (S_{database}).

In pratica: Questo meccanismo incentiva l'Agente a "giocare" con il campione, a esplorare macchie apparentemente insignificanti o a tentare combinazioni di parametri inusuali, sapendo che se trova uno spettro chimicamente *nuovo*, riceverà un premio per la sua audacia esplorativa. In questo modo, il sistema cerca attivamente l'ignoto.