Reinforcement Learning-driven Information Seeking: A Quantum Probabilistic Approach

Amit Kumar Jaiswal[1] Haiming Liu[1] Ingo Frommholz[1]

[1] Institute for Research in Applicable Computing,
University of Bedfordshire, United Kingdom

University of Bedfordshire

QUARTZ Quantum Information Access and Retrieval Theory

July 30, 2020

Outline

Background

- Uncertainty in information seeking (IS) decreases as the user proceeds through the seeking process [Chowdhury 2011].
- In general, user interaction keep searchers (in an **information seeking process as Foraging** [Wittek 2016, Pirolli 1999]) to not consume the information goal.
- **Reinforcement learning** to generalise user (or forager) search behaviour by their **action representation and transformation** [Chandak 2019].
- Simultaneous happening of risk and ambiguity (**aspect of uncertainty**) during the IS process in turn converges with the **quantum theory** [Wittek 2016, Piwowarski 2010]

What's new here?

- To guide searcher (or forager) during information seeking process (especially information exploration) by means of **Reinforced Foraging** mechanism.
  - **Reinforced Foraging:** Reinforcement learning help's us devise the Information Foraging strategy to follow the feat of information seeking.
  - **Assumption:** We consider uncertainty in IS to be a problem that is closely related to information need.
- Representation of user actions (i.e. queries as information need) follows the **quantum probabilistic constructs** [Van Rijsbergen 2004].
- Theoretical framework that describes **guided information seeking** powered by quantum-parameterised reinforced foraging.

Trivia

QUARTZ
Quantum Information Access and Retrieval Theory

Why RL?

- There is no supervision, only a reward signal.
- Feedback is delayed, not instantaneous.
- Agent's actions effect subsequent data it receives.

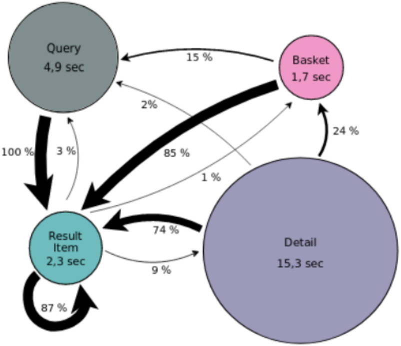
Central idea of RL:

- Interacts with the environment.
- Learns from experience.
- The target is to get the maximum expected cumulative rewards.

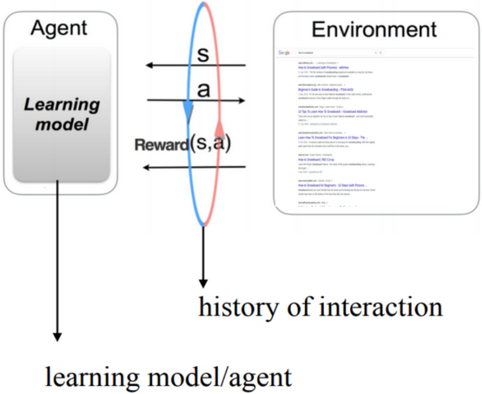
Central idea of Information Foraging theory (IFT):

- Searches via information patches and constantly makes decision among it.
- Learns from enrichment.
- The target is to get as much relevant information in as little time as possible.

Markov Decision Process for RL



Markov chain for User search process in IIR (Tran 2012)



RL: Agents learn behaviour not data

Figure: Recall Markov Property

"The future is independent of the past given the present"

- An information environment for reinforcement learning follows Markov decision process.
- In RL, the agent **changes the environment.**

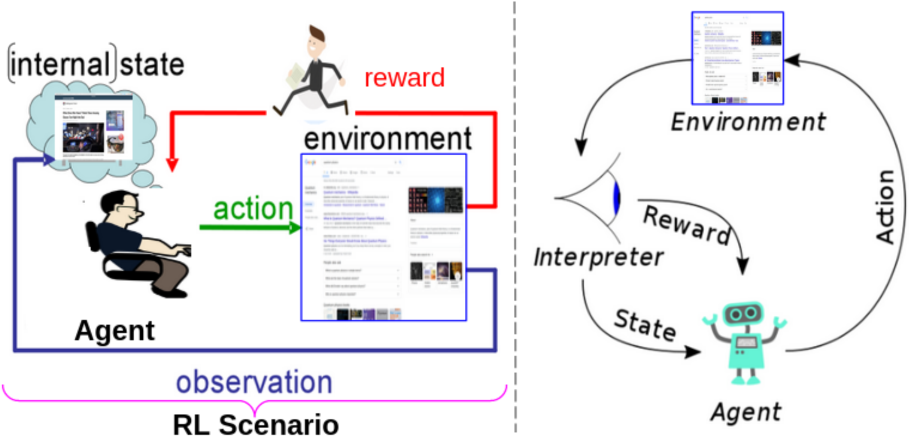


Figure: From left to right: Scenario of Reinforcement Learning and its process

Information Foraging Theory

- A theory of **human information seeking**, eventually derived from optimal foraging theory [Pirolli 1999].
- Adaptive process with regard to optimal use of knowledge about expected information value, expected cost of acquiring relevant information.
- Activities associated with seeking, acquiring, and dealing information sources

An Example: Spiders/Wolves

- IFT comprises of three main constructs:
  - Information Patch
  - Information Scent
  - Information Diet

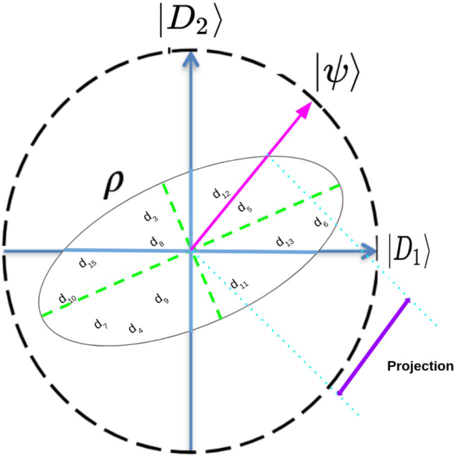
RL with IFT: Reinforced Foraging

- Hypothesis: **Information seeker as Forager** [Wittek 2016] as **RL agent**.
- Seeker adopts foraging behaviour (**explore as well as exploit**).
- Reinforcement learning process enhanced by such type of information seekers - so called, an **adaptive RL agent.**
- IFT can resolve RL limitation of **delayed reward** i.e. "why every step of seeker is important".
- Foraging behaviour can enhance "experience" in reinforcement learning mechanism.

Apply to Information Seeking

- RL agent "interact with the environment", whereas in IS "the process of seeking may provide the learning required to satisfy one's information need".
- "Learn by doing with delayed reward" aspect of RL when meets IFT lead to IS process.
- IFT supports RL in **"exploration & exploitation"**.
  - Available actions as "exploration".
  - Positively rewarded actions are drawn as "exploitation".
- IS behaviour pattern when meets this trade-off makes **user actions uncertain** [Wittek 2016].
  - Such behaviour in IFT is mostly sequential.
  - This type of uncertainty can be described using quantum theory.
- IFT can resolve RL limitation of **delayed reward** i.e. "why every step of seeker is important".
- Foraging behaviour can enhance "experience" in reinforcement learning mechanism.

Quantum Probability Theory: Probability theory based on Hilbert space formalism



Constructs of quantum-inspired RL framework

QUARTZ
Quantum Information Access and Retrieval Theory

Agent : In our framework, the agent is a forager (information seeker).

Action : The agent executes query (as action $|a_t\rangle$, receives states ($|s_t\rangle$) and a scalar reward ($R_{e_{i,a_i}}$).

Environment : Receives agent action (query) and emits observation ($|s_{t+1}\rangle$ with corresponding reward.

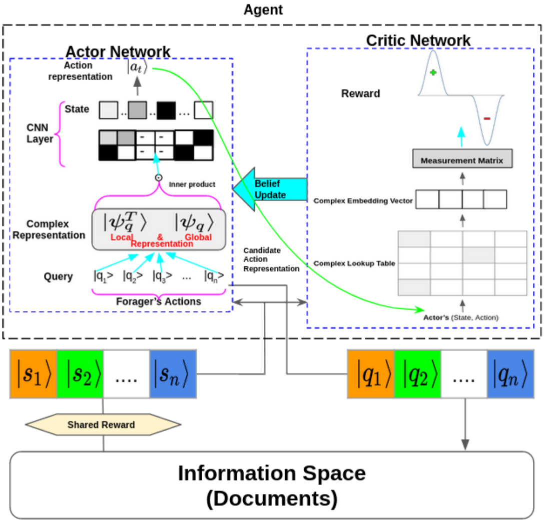
State : In our case, a state can be seen as the product of the probability amplitudes of global-local projection (word meanings) for all words of a query.

State Transition : We use a feedback mechanism to compute the transition among the states.

Policy : We use stochastic policy network, so called Actor-Critic reinforcement learning method [Lowe 2017].

Reward : The success value of an agent's action ($|q_i\rangle$)

Quantum-inspired Reinforcement Learning Framework



Contributions

Quantum-like reinforcement learning framework that incorporates Information Foraging strategy for information seeking to

- model the information foragers' behaviour, where the Actor-critic method to enhance the agent's experience in a text query-matching task.
- learn the policy where query representation is parameterised using quantum language models.

Conclusion

- An initial attempt to encapsulate the foraging behaviour in a principled RL framework.
- Characterising information seeking in a formal behavioural model that delineates uncertainty in users' information need.

References I

Yash Chandak, Georgios Theocharous, James Kostas, Scott Jordan et Philip Thomas.
Learning Action Representations for Reinforcement Learning.
In International Conference on Machine Learning, pages 941–950, 2019.

Sudatta Chowdhury, Forbes Gibb et Monica Landoni.
Uncertainty in information seeking and retrieval: A study in an academic environment.
Information Processing & Management, vol. 47, no. 2, pages 157–175, 2011.

References II

Ryan Lowe, Yi I Wu, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel et Igor Mordatch.
Multi-agent actor-critic for mixed cooperative-competitive environments.
In Advances in neural information processing systems, pages 6379–6390, 2017.

Peter Pirolli et Stuart Card.
Information foraging.
Psychological review, vol. 106, no. 4, page 643, 1999.

Benjamin Piwowarski, Ingo Frommholz, Mounia Lalmas et Keith Van Rijsbergen.
What can quantum theory bring to information retrieval.
In Proceedings of the 19th ACM international conference on Information and knowledge management, pages 59–68, 2010.

References III

Cornelis Joost Van Rijsbergen.
The geometry of information retrieval.
Cambridge University Press, 2004.

Peter Wittek, Ying-Hsang Liu, Sándor Darányi, Tom Gedeon et Ik Soo Lim.
Risk and ambiguity in information seeking: Eye gaze patterns reveal contextual behavior in dealing with uncertainty.
Frontiers in psychology, vol. 7, page 1790, 2016.

Reinforcement Learning Setup

Notation	Interpretation	Description
α_{i,b_i}	$b_i \in \{1, ..., k\}$	Probability amplitude
$\lvert\phi_{b_i}\rangle$	Semantic meaning	Basis vector (n (word vectors) or k^n dimension for tensor product of basis vectors)
$\lvert w_i\rangle$	$\sum_{b_i=1}^{k} \alpha_{i,b_i} \lvert\phi_{b_i}\rangle$	Word state vector
$\lvert q_i\rangle$	$\lvert w_1\rangle \otimes \lvert w_2\rangle ... \otimes \lvert w_n\rangle$	Query state vector
$\lvert\psi_q^T\rangle$	$\sum_{b_1,...,b_n=1}^{k} (\underbrace{\prod_{i=1}^{n} \alpha_{i,b_i}}_{\mathcal{L}_{b_1,...,b_n}} \lvert\phi_{b_i}\rangle \otimes ... \otimes \lvert\phi_{b_n}\rangle)$	Local representation ($\mathcal{L}$ is a k^n dimensional tensor)
$\lvert\psi_q\rangle$	$\sum_{b_1,...,b_n=1}^{k} \mathcal{G}_{b_1 b_2 ... b_n} \lvert\phi_{b_i}\rangle \otimes ... \otimes \lvert\phi_{b_n}\rangle$	Global representation of combined meanings/patches
$\mathcal{G}$	$\sum_{r=1}^{R} w_r \cdot e_{r,1} \otimes e_{r,2} \otimes ... \otimes e_{r,n}$	Probability amplitude (semantic space of meaning)
$\langle\psi_q^T \vert \psi_q\rangle$	$\sum_{b_1,...,b_n=1}^{k} \underbrace{\mathcal{G}_{b_1...b_n} \times \prod_{i=1}^{n} \alpha_{i,b_i}}_{Probability\ amplitudes}$	Projection of the global representation to the local representation of a query
State	$\prod_{i=1}^{n} \sum_{b_i=1}^{k} e_{r,i,b_i} \cdot \alpha_{i,b_i}$	Actor network state module (product pooling layer [35])
$\lvert a_t\rangle$	$(\lvert a_1\rangle, \lvert a_2\rangle, ..., \lvert a_R\rangle)^T$	Output of the Actor network