

The Quantum Probability Ranking Principle

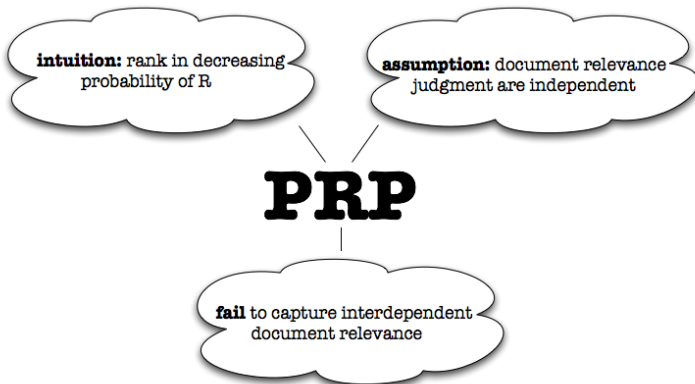
ICTIR 2009

Guido Zuccon, Leif Azzopardi, Keith van Rijsbergen

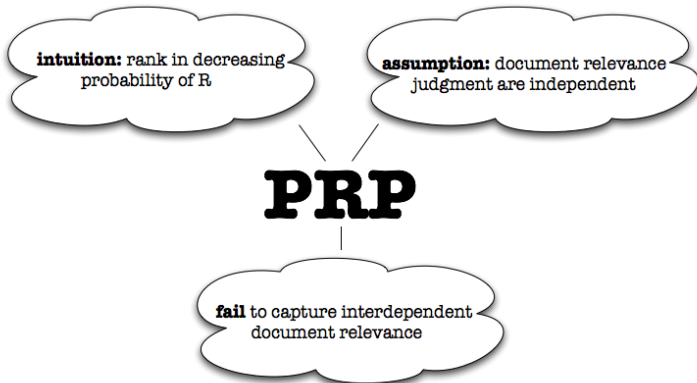
University of Glasgow
guido@dcs.gla.ac.uk



University
of Glasgow | Department of
Computing Science

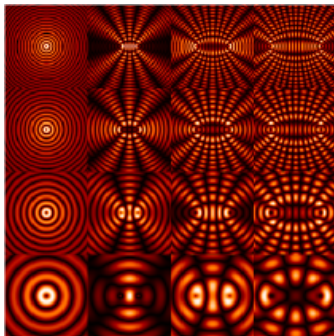


goal: model interdependent document relevance in ranking process.



goal: model interdependent document relevance in ranking process.

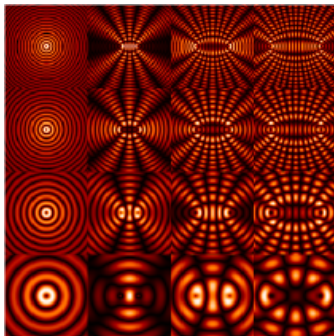
intuition: quantum probability theory accounts for interference
between events



(from wikipedia.org)

program: use interference in document ranking to model
interdependencies

intuition: quantum probability theory accounts for interference between events

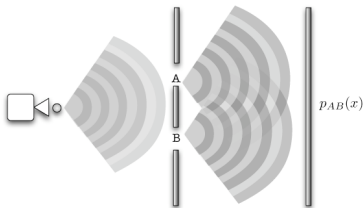


(from wikipedia.org)

program: use interference in document ranking to model interdependencies

How much is $p_{AB}(x)$?

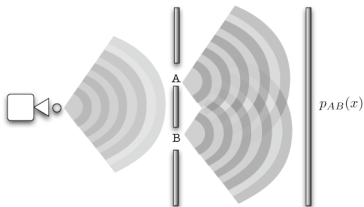
- $p_A(x)$: Pr. hitting when B is close (i.e. passing through A).
- $p_B(x)$: Pr. hitting when A is close (i.e. passing through B).



- Kolmogorovian law of total probability: $p_{AB}(x) = p_A(x) + p_B(x)$
- Quantum law of total probability:
 $p_{AB}(x) = p_A(x) + p_B(x) + I_{AB}(x)$

How much is $p_{AB}(x)$?

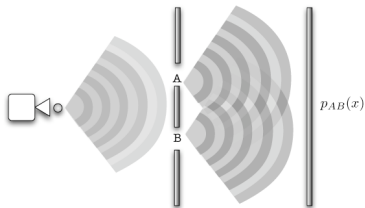
- $p_A(x)$: Pr. hitting when B is close (i.e. passing through A).
- $p_B(x)$: Pr. hitting when A is close (i.e. passing through B).



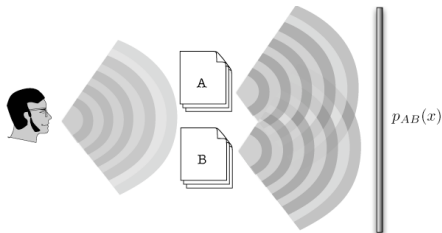
- Kolmogorovian law of total probability: $\mathbf{p}_{AB}(\mathbf{x}) = \mathbf{p}_A(\mathbf{x}) + \mathbf{p}_B(\mathbf{x})$
- Quantum law of total probability:
 $\mathbf{p}_{AB}(\mathbf{x}) = \mathbf{p}_A(\mathbf{x}) + \mathbf{p}_B(\mathbf{x}) + \mathbf{I}_{AB}(\mathbf{x})$

How much is $p_{AB}(x)$?

- $p_A(x)$: Pr. hitting when B is close (i.e. passing through A).
- $p_B(x)$: Pr. hitting when A is close (i.e. passing through B).



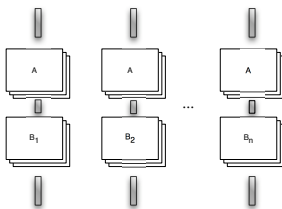
- Kolmogorovian law of total probability: $\mathbf{p}_{AB}(\mathbf{x}) = \mathbf{p}_A(\mathbf{x}) + \mathbf{p}_B(\mathbf{x})$
- Quantum law of total probability:
 $\mathbf{p}_{AB}(\mathbf{x}) = \mathbf{p}_A(\mathbf{x}) + \mathbf{p}_B(\mathbf{x}) + \mathbf{I}_{AB}(\mathbf{x})$



How much is user's satisfaction for ranking A, B ($p_{AB}(x)$)?

- Kolmogorovian law of total probability: $\mathbf{p}_{AB}(\mathbf{x}) = \mathbf{p}_A(\mathbf{x}) + \mathbf{p}_B(\mathbf{x})$
- Quantum law of total probability:
 $\mathbf{p}_{AB}(\mathbf{x}) = \mathbf{p}_A(\mathbf{x}) + \mathbf{p}_B(\mathbf{x}) + \mathbf{I}_{AB}(\mathbf{x})$

- scenario:** several screens are available, all of them have in common slit A but each screen i is characterized by a slit B_i



- task:** maximize expected utility of presented documents \implies
 maximize $p_{AB_i}(x) \implies$ find pair (A, B_i) which maximizes $p_{AB_i}(x)$

Choosing the document to rank: $\operatorname{argmax}_{B_i}(p_{AB_i}(x)) = \dots$

- Kolmogorovian: $\operatorname{argmax}_{B_i}(p_{B_i}(x)) \Leftarrow$ **PRP**
- Quantum: $\operatorname{argmax}_{B_i}(p_{B_i}(x) + I_{AB_i}(x)) \Leftarrow$ **QPRP**

Interference has to be accounted for: QPRP

Choice of document to select among not only depends upon $p_{B_i}(x)$ but also on interference, $I_{AB_i}(x)$.

- The **first document** to rank is the one with **highest** $p_A(x)$
- B should be ranked after A and before C i.f.f.
 $p_B(x) + I_{AB}(x) \geq p_C(x) + I_{AC}(x)$: **account for interference**

Choosing the document to rank: $\operatorname{argmax}_{B_i}(p_{AB_i}(x)) = \dots$

- Kolmogorovian: $\operatorname{argmax}_{B_i}(p_{B_i}(x)) \Leftarrow$ **PRP**
- Quantum: $\operatorname{argmax}_{B_i}(p_{B_i}(x) + I_{AB_i}(x)) \Leftarrow$ **QPRP**

Interference has to be accounted for: QPRP

Choice of document to select among not only depends upon $p_{B_i}(x)$ but also on interference, $I_{AB_i}(x)$.

- The **first document** to rank is the one with **highest** $p_A(x)$
- B should be ranked after A and before C i.f.f.
 $p_B(x) + I_{AB}(x) \geq p_C(x) + I_{AC}(x)$: **account for interference**

When does B have to be ranked above C ?

	$p_B(x) > p_C(x)$	$p_B(x) = p_C(x)$
PRP	B before C	either
QPRP	B before C iff $p_B(x) - p_C(x) > I_{AC}(x) - I_{AB}(x)$	B before C iff $I_{AB}(x) > I_{AC}(x)$

Matter of empirical investigation determine how many times the rankings actually differ.

Interference term in details

$$I_{AB}(x) = 2\sqrt{p_A(x)}\sqrt{p_B(x)}\cos\theta_{AB}$$

Behavior of I_{AB} is governed by θ_{AB} .

- Phases actively affect document ranking.
- e.g., when $p_B(x) \geq p_C(x)$ interference capable of subverting ordering of PRP if

$$\frac{p_B(x) - p_C(x)}{2\sqrt{p_A(x)}} < \sqrt{p_C(x)} \cos \theta_{AC} - \sqrt{p_B(x)} \cos \theta_{AB}$$

- $p_A(x)$, $p_B(x)$, etc, are usually estimated from statistical features of the document collection.
- what about θ ? Still an open question
- naïve idea: de-phasing arccosine similarity/ (–) Pearson's correlation

- 1 novel application of QT in IR
- 2 QPRP models **interdependent document relevance**
- 3 **interference** is the key

To do/open questions:

- how to **estimate** interference? (or initial complex amplitude distribution)
- empirical investigation on ranking **differences** b/w PRP and QPRP
- **application** to IR tasks
- interdependent relevance is modeled in principled way: do strategies for IDR uphold for the QPRP (and not in PRP), e.g. [Chen and Karger, 2006, Wang and Zhu, 2009, Zhai et al., 2003] models.

- 1 novel application of QT in IR
- 2 QPRP models **interdependent document relevance**
- 3 **interference** is the key

To do/open questions:

- how to **estimate** interference? (or initial complex amplitude distribution)
- empirical investigation on ranking **differences** b/w PRP and QPRP
- **application** to IR tasks
- interdependent relevance is modeled in principled way: do strategies for IDR uphold for the QPRP (and not in PRP), e.g. [Chen and Karger, 2006, Wang and Zhu, 2009, Zhai et al., 2003] models.

[Chen and Karger, 2006] Chen, H. and Karger, D. R. (2006).

Less is more: probabilistic models for retrieving fewer relevant documents.

In SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, pages 429–436, New York, NY, USA. ACM.

[Wang and Zhu, 2009] Wang, J. and Zhu, J. (2009).

Portfolio theory of information retrieval.

In SIGIR '09: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pages 115–122, New York, NY, USA. ACM.

[Zhai et al., 2003] Zhai, C. X., Cohen, W. W., and Lafferty, J. (2003).

Beyond independent relevance: methods and evaluation metrics for subtopic retrieval.

In SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, pages 10–17, New York, NY, USA. ACM.

Questions?