A Quantum-Cognitive Perspective for Information Interaction and Retrieval

Dawei Song
Beijing Institute of Technology

dawei.song2010@gmail.com

Acknowledgement

The speaker thanks the following collaborators for their contributions:

Peter Bruza (Queensland University of Technology, Australia),

Jerome Busemeyer (Indiana University, United States),

Sagar Uprety (The Open University, United Kingdom),

Qiuchi Li and Benyou Wang (Padua University, Italy),

Peng Zhang and Yazhou Zhang (Tianjin University, China)

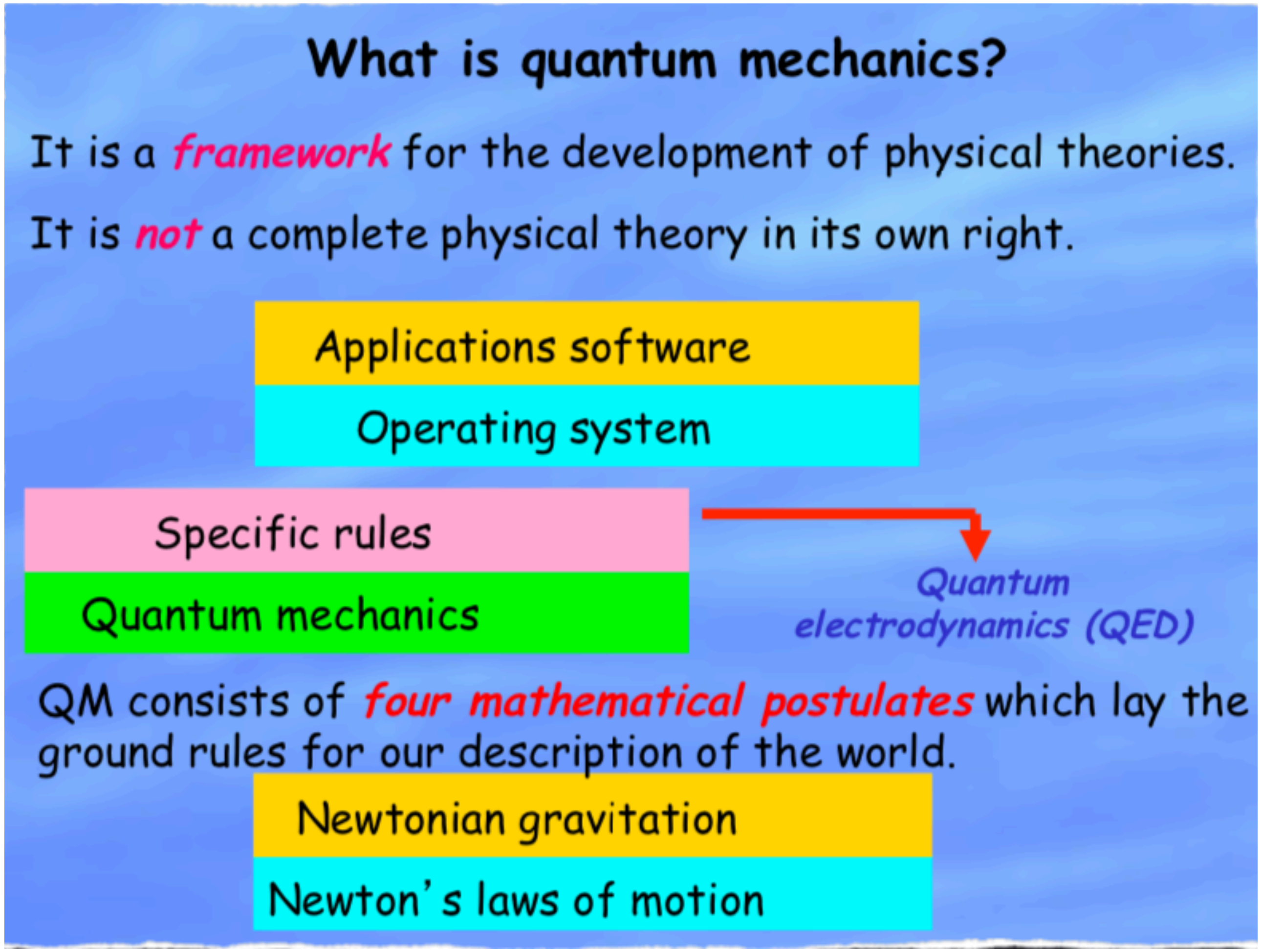
Information Retrieval

- IR is the process of finding desired information that is relevant to a user's information need.
- By nature IR is a cognitively situated task involving user's continuous interaction, learning and decision making about information.
  - Relevance
  - Text understanding

Quantum inspired IR models

- Pioneering work by Keith van Rijsbergen

van Rijsbergen “Quantum Haystacks”, Salton Award Lecture, SIGIR2006



"one has a formal mechanism in which logics and probability theory arise simultaneously and are derived simultaneously"

- J. von Neumann

van Rijsbergen's view

Geometry:

Quantum probability and logic

Subspace logic - esp. negation
Probability via Gleason
QP is a generalisation of classical probability
Closure is a linear span/combination
Logic + probability + distance + measurement
Perspective/point of view

- Logical, vector space, probabilistic and language models can be formulated in a unified framework of QT
- Retrieval from incompatible perspectives (Principle of Uncertainty)

Development of Quantum IR formal models

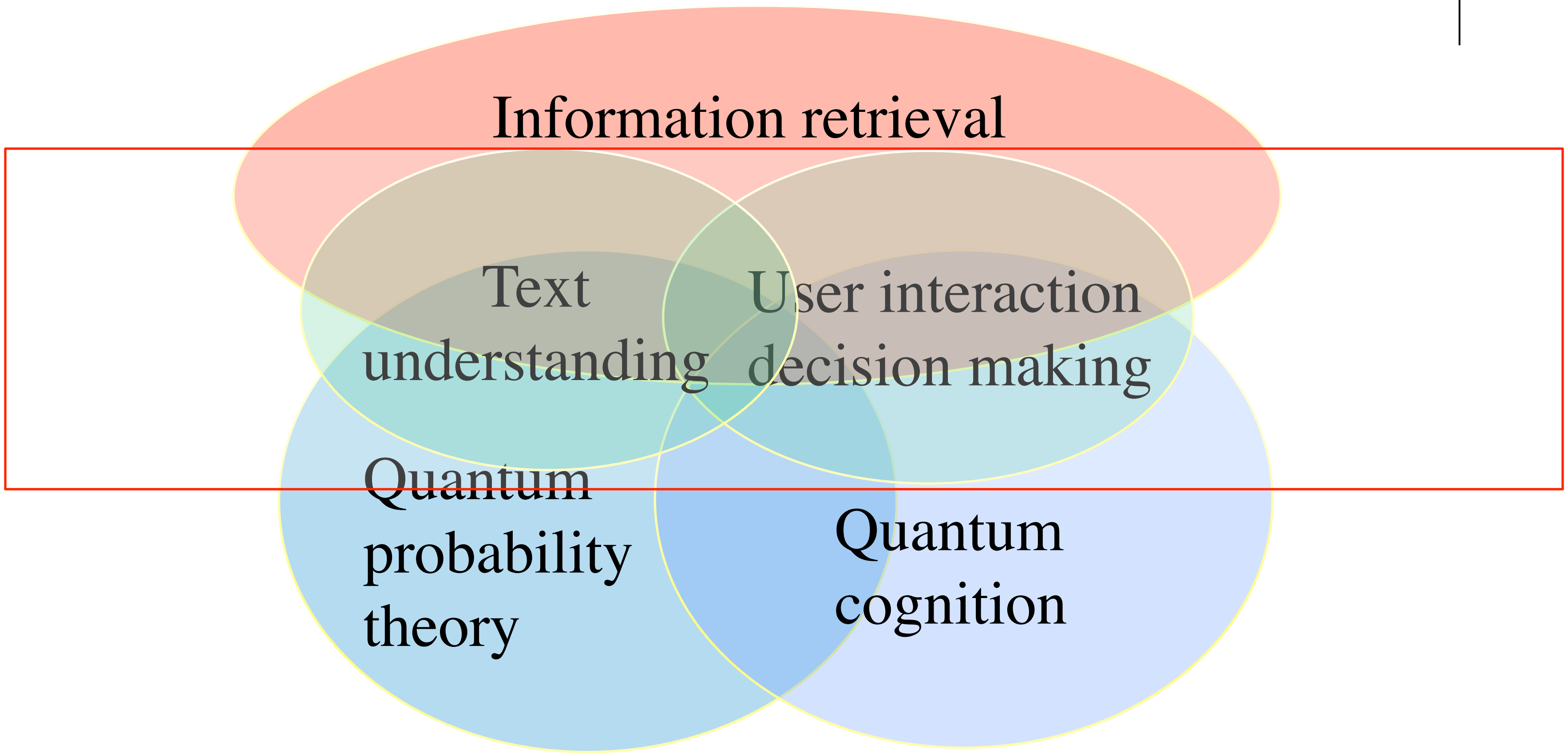
Milestones

Quantum Analogy based IR Methods	Quantum Language Models (QLMs)	Quantum Language Models in Neural Network structures
Double Slit	*Original QLM*	*NN-based QLM*
Photon Polarization	*Adaptive QLM*	*Tensor space QLM*
Quantum measurement		

What's missing here:

Are these simply a try-out of another apparent relevant theory in IR? Is there a fundamental quantum-like structure in IR and in what aspects, so as to the QM framework is necessary?

A Quantum Cognitive View of IR



- "quantumeness" of users in relevance decision making
- Text representation in line with human understanding of text

Quantum Cognition

- Quantum Theory – generalised theory of probability.
- Human decision-making under uncertainty is more quantum-like rather than classical.
- Quantum Cognition – offers an alternative way to build probabilistic models for human decision-making under uncertainty.
- Utilizes mathematical tools of Quantum Theory.
  - Complex Hilbert space
  - Superposition
  - Projective measurement over subspaces
  - $P_A P_B \neq P_B P_A$ - incompatible decision perspectives
  - Contextuality

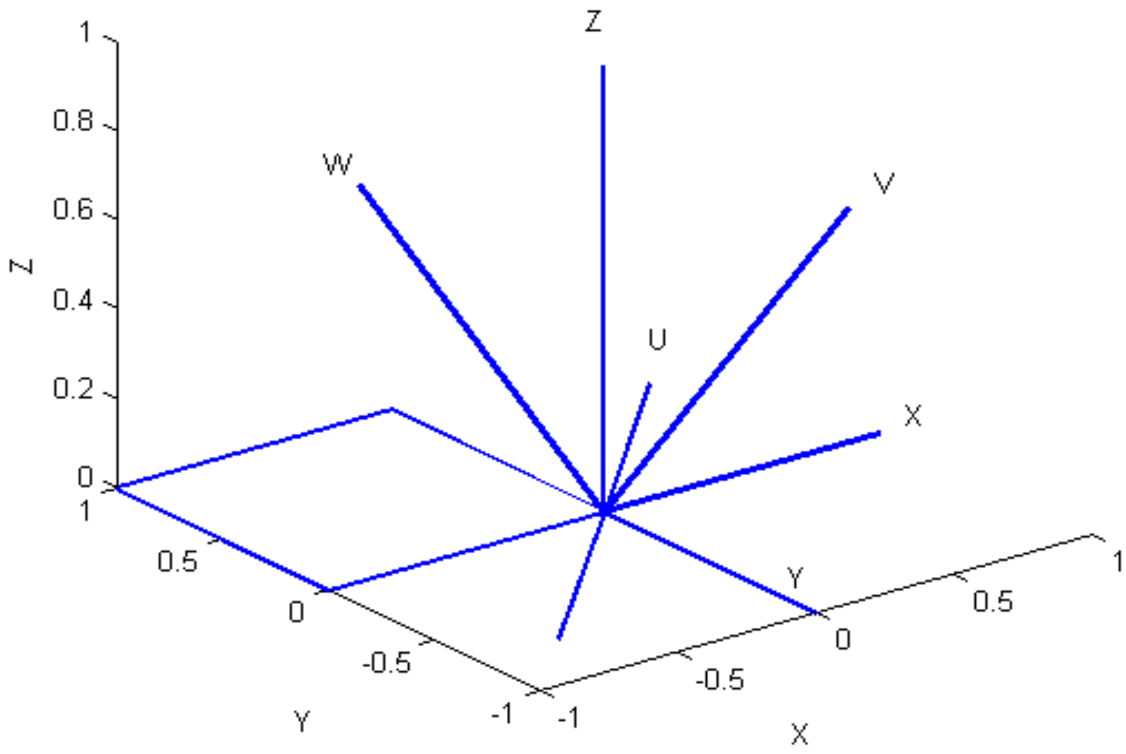
Events

- Classic theory
  - Universal set contains all elements
  - Events are subsets of a universal set
- Quantum theory
  - Vector space spanned by all eigenvectors
  - Events are subspaces of a vector space

Example

- First set of questions — Event
  - Will you vote democrat? — X
  - Will you vote republican? — Y
  - Will you vote independent? — Z
- Second set of questions
  - Are you a moderate? — U
  - Are you a liberal? — V
  - Are you a conservative? — W

A Vector Space Representation



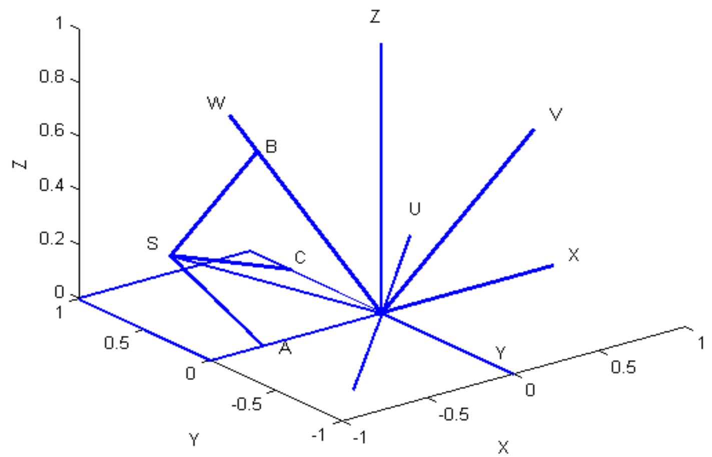
Eigenvector Set 1:
X=democrat
Y=republican
Z=independent

Eigenvector Set 2:
U=moderate
V=liberal
W= conservative

State of the System

- Classic probability
  - A probability function ***p*** assigns probabilities to events (subsets)
- Quantum probability
  - A state vector $|\psi\rangle$ assigns probabilities to quantum events (subspaces), as density matrices

Projectors



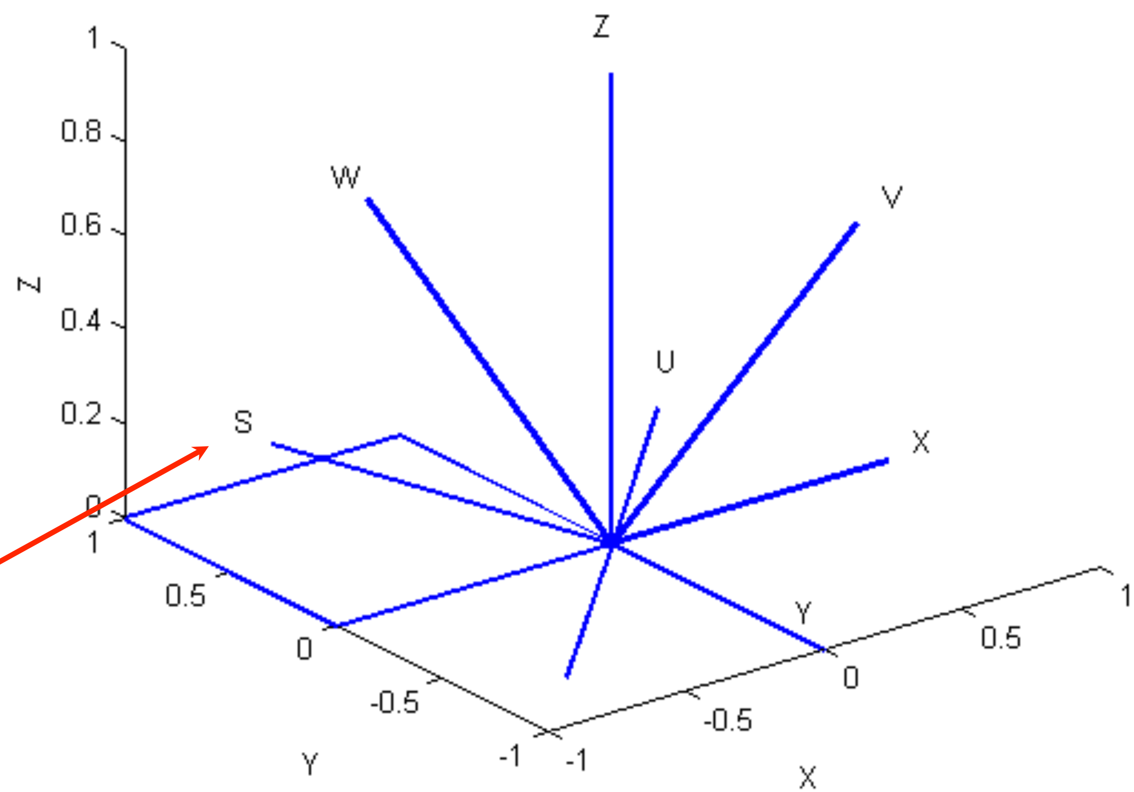
$$X = \begin{bmatrix}1\\0\\0\end{bmatrix}, Y = \begin{bmatrix}0\\1\\0\end{bmatrix}, Z = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

$$M_X = X \cdot X^\dagger = \begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}1 & 0 & 0\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}$$

$$M_X = X \cdot X^\dagger = \begin{bmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix},\ M_Y = Y \cdot Y^\dagger = \begin{bmatrix}0 & 0 & 0\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix},\ M_Z = Z \cdot Z^\dagger = \begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 1\end{bmatrix}$$

$$M_X + M_Y + M_Z = I$$

Quantum State

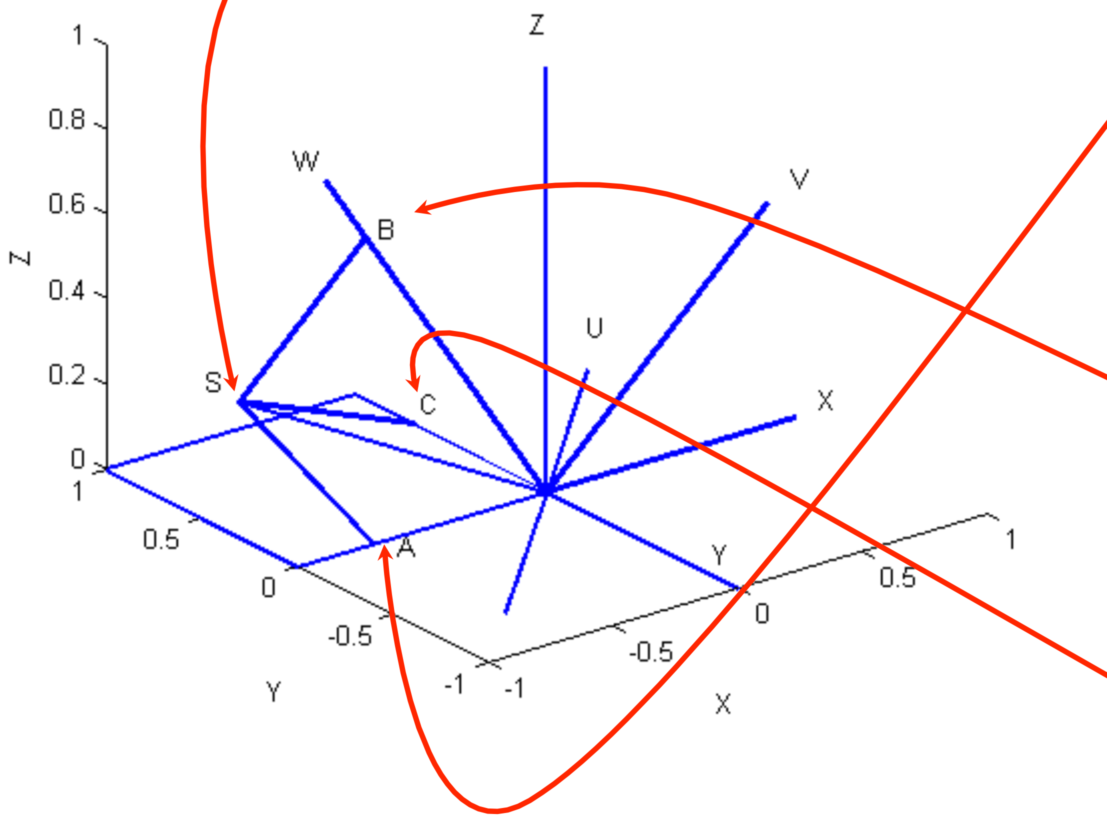


$|\psi\rangle$ = S = State

$|\psi\rangle$=S = (-.6963)X + (0.6963)Y + (0.1741)Z

$|\psi\rangle$=S = (0.000)U + (-0.5732)V + (0.8194)W

Quantum State $|\psi\rangle$



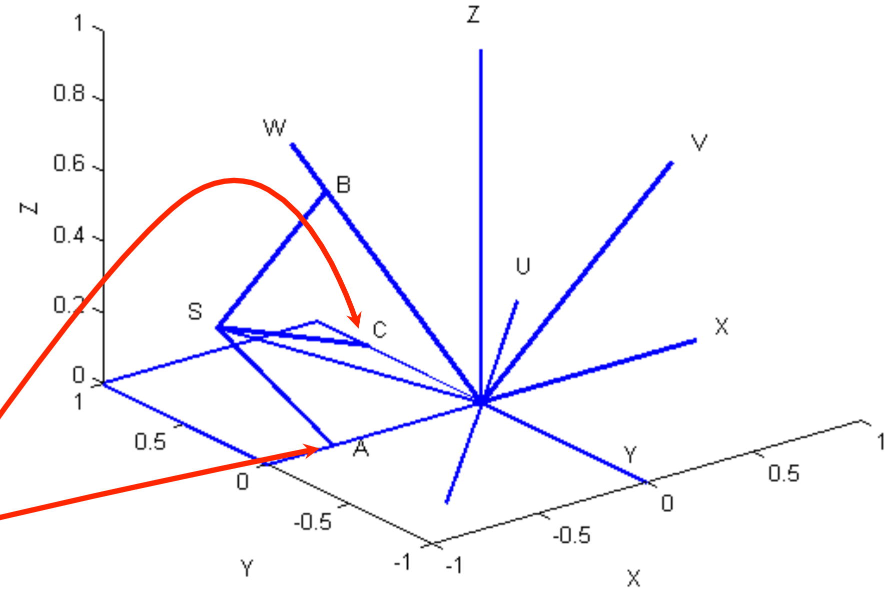
$|\psi\rangle$ = S = State

$M_X S$ = A = projection on X
$Q(X) = |A|^2$ (pr. vote democrat)

$M_W S$ = B = projection on W
$Q(W) = |B|^2$ (pr. conservative)

$M_Y S$ = C = projection on Y
$Q(Y) = |C|^2$ (pr. vote republican)

$|\psi\rangle = S = (-.6963)X + (0.6963)Y + (.1741)Z$
X,Y,Z are canonical basis vectors



$$A = M_X \cdot S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot S = \begin{bmatrix} -.6963 \\ 0 \\ 0 \end{bmatrix}$$

$Q(X)=|A|^2 = |-.6963|^2 = .4848$

(pr. vote democrat)

$$C = M_Y \cdot S = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot S = \begin{bmatrix} 0 \\ 0.6963 \\ 0 \end{bmatrix}$$

$Q(Y)=|C|^2 = |+.6963|^2 = .4848$

(pr. vote republican)

Compatibility

- Quantum theory allows for two kinds of events
  - Compatible events are described by a common set of eigenvectors
  - Incompatible events are not described by a common set of eigenvectors
- Classic probability theory essentially assumes all events are described by a common set of elements --in other words, compatible.

Test for Compatibility

- Incompatible events

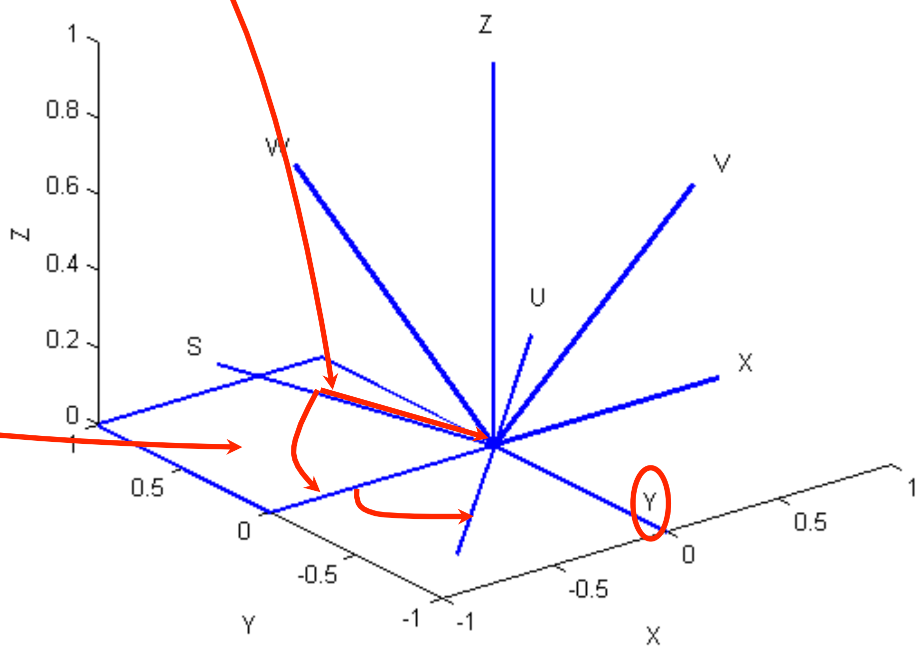
  (order matters)

  $M_U \cdot M_X \neq M_X \cdot M_U$

- Compatible Events

  (order does not matter)

  $M_X \cdot M_Y = M_Y \cdot M_X$



Probability of Conjunction (A and B)

$$Q(A)Q(B|A) = |M_A \cdot \psi|^2 \cdot |M_B \cdot \psi_A|^2$$

$$= |M_A \cdot \psi|^2 \cdot \left| M_B \cdot \frac{M_A \cdot \psi}{|M_A \cdot \psi|} \right|^2$$

$$= |M_B \cdot M_A \cdot \psi|^2$$

Depends on order for incompatible events:

$M_A \cdot M_B \neq M_B \cdot M_A$

Algebra of Events

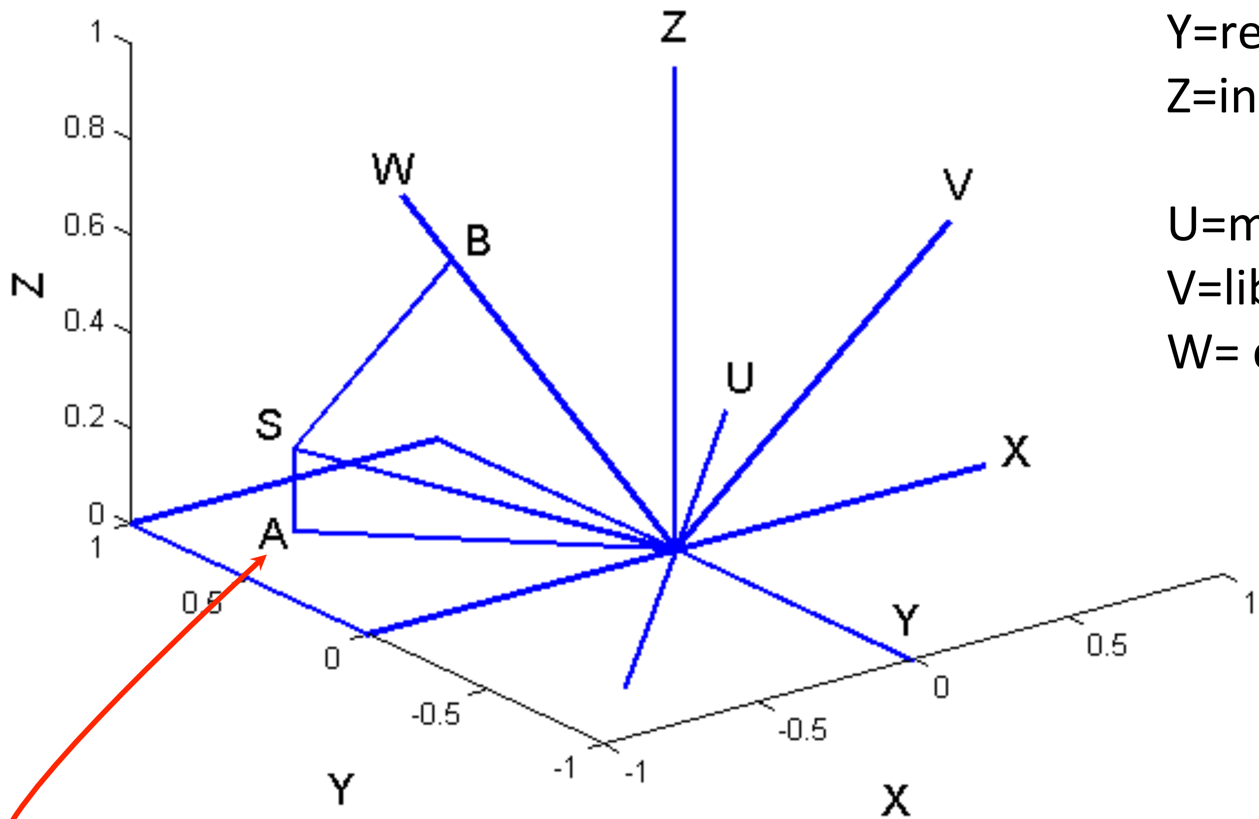
- Classic probability
  - Events form a Boolean Algebra
  - Commutative and Distributive axioms required
  - Law of total probability obeyed
- Quantum probability
  - Events form a Partial Boolean Algebra
  - Commutative and Distributive axioms not necessary
  - Law of total probability violated

Probability of X or Y

$|\psi\rangle$ = S = state of voter

X=democrat
Y=republican
Z=independent

U=moderate
V=liberal
W= conservative



$$M_{X+Y} = M_X + M_Y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$M_{X+Y} \cdot S = \begin{bmatrix} -.6963 \\ .6969 \\ 0 \end{bmatrix}$$

$$Q(X \text{ or } Y) = |M_{X+Y} \cdot S|^2 = |-.6963|^2 + |.6969|^2 = .9697$$

(pr. vote democrat or republican)

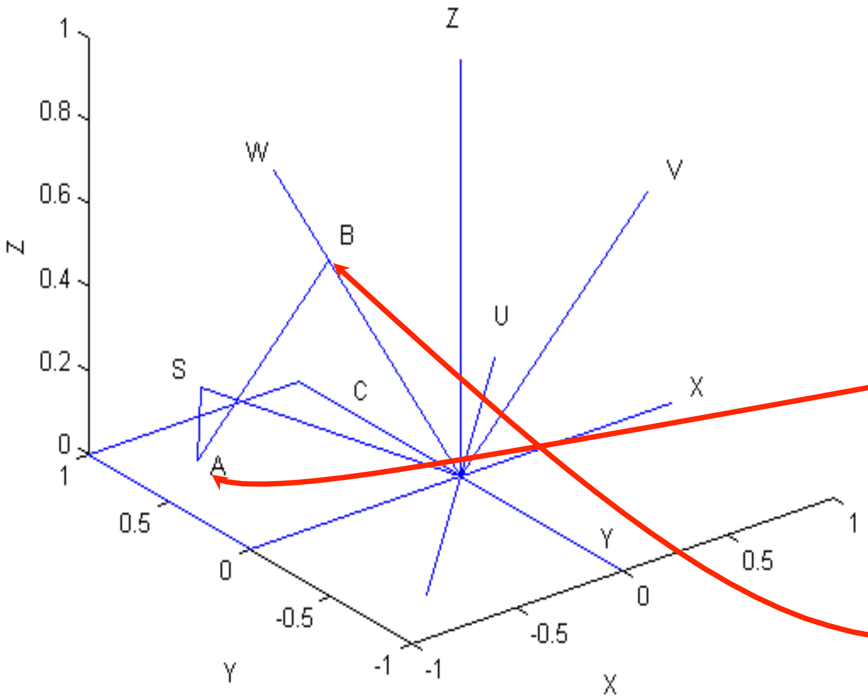
Complementary Event ~(X or Y)

$$M_Z = (I - M_{X+Y}) = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} - \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 1\end{bmatrix}$$

$$M_Z \cdot S = (I - M_{X+Y}) \cdot S = \begin{bmatrix}0\\0\\.1741\end{bmatrix}$$

$$\begin{aligned} Q(\sim (X \text{ or } Y)) &= |(I - M_{X+Y}) \cdot S|^2 \\ &= Q(Z) = |M_Z \cdot S|^2 = |.1741|^2 = .0303 \\ &= 1 - Q(X \text{ or } Y) \end{aligned}$$

Lüder's Rule: Conditional Probabilities



1. Project state
   $A = M_{X+Y} \cdot S$

2. Normalize
   $\psi_A = A/|A|$

3. Project to new state
   $M_W \cdot \psi_A$

4. Conditional Probability
   $Q(W|X \text{ or } Y) = |M_W \cdot \psi_A|^2$

$Q(W|\ X \text{ or } Y) = .52$

(pr. Conservative given vote democrat or republican)

Probability of Conjunction (A and B)

$$Q(A)Q(B|A) = |M_A \cdot \psi|^2 \cdot |M_B \cdot \psi_A|^2$$

$$= |M_A \cdot \psi|^2 \cdot \left| M_B \cdot \frac{M_A \cdot \psi}{|M_A \cdot \psi|} \right|^2$$

$$= |M_B \cdot M_A \cdot \psi|^2$$

Depends on order for incompatible events:

$M_A \cdot M_B \neq M_B \cdot M_A$

Violation of Commutative Property

X and then U

$Q(X)Q(U|X)$

$= |M_U \cdot M_X \cdot \psi|^2$

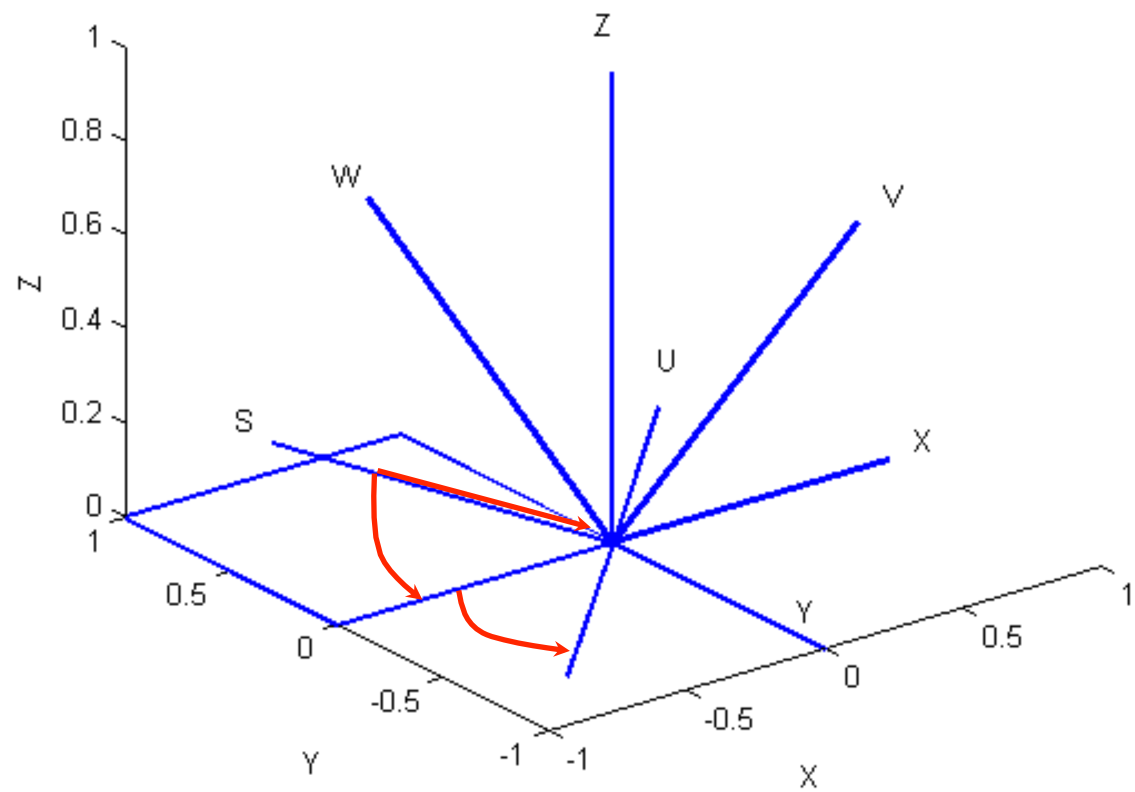
$= .2424$

U and then X

$Q(U)Q(X|U)$

$= |M_X \cdot M_U \cdot \psi|^2$

$= 0$



Violation of Law of total probability

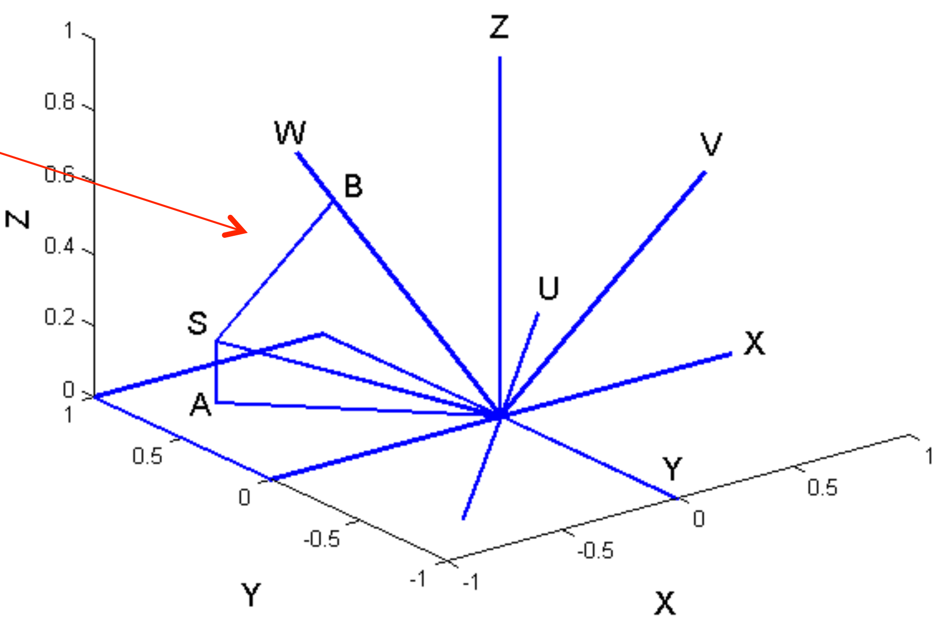
Q(Z) = .0303 and Q(~Z) = .9797

Q(W|Z) = .50 and Q(W|~Z) = .52

Total Prob: Q(W) = Q(Z and W)+ Q(~Z and W) = Q(Z)Q(W|Z) + Q(~Z)Q(W|~Z) = .9848

But Q(W) = .6714 = Q(Z)Q(W|Z) + Q(~Z)Q(W|~Z) + Int

$$\text{Int} = 2\cdot \mathrm{Re}[(M_W \cdot M_Z\, \psi)^T \cdot (M_W \cdot M_{\sim Z}\, \psi)]$$



How does it have to do with user cognition and decision making?

Five Principles that challenge the Current Foundations of Cognitive Science

1. Cognitive Measures Create rather than Record

→ We construct new judgments rather than simply record existing judgments about complex emotional events

- Quantum measures not always pre-defined
- Constructed at the point of interaction with information

Classic Information Processing

Cognitive System is in a definite state with respect to each possible measure

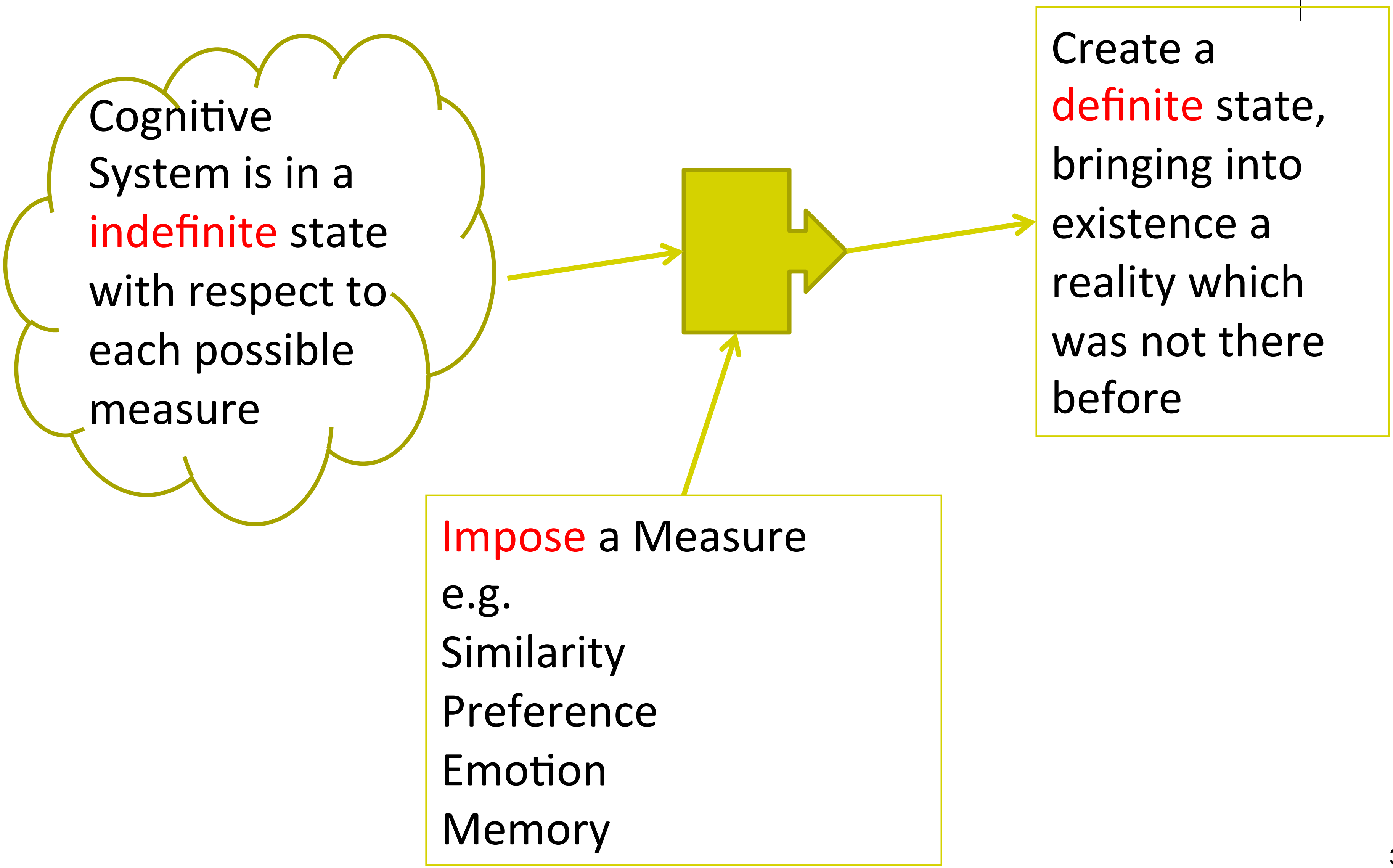
→

Take a Measure pre-defined e.g.
Similarity
Preference
Emotion
Memory

→

Simply record what existed immediately before our Measurement was taken

Quantum Information Processing

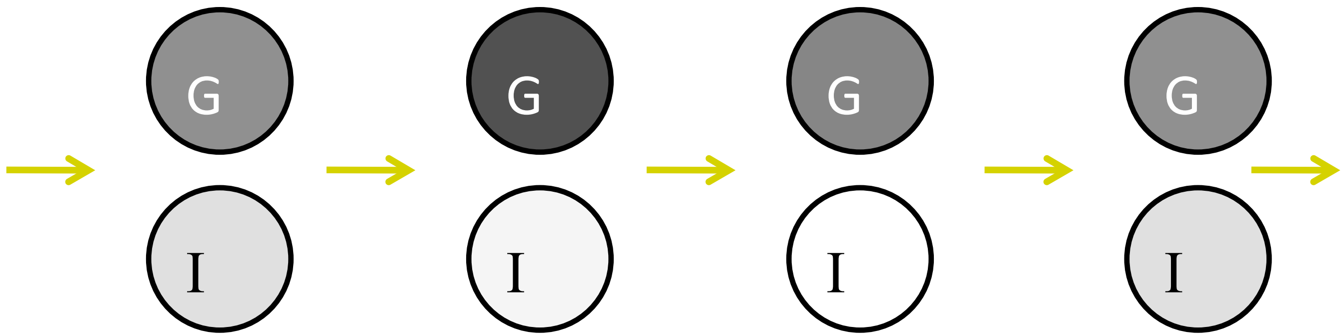


2. Cognition behaves like a wave rather than a particle

Classic Information Processing

Is the defendant Guilty or Innocent?

Quantum Information Processing



Is the defendant Guilty or Innocent?

→ Our beliefs are superposed – we don't jump from state to state, instead we experience a feeling of ambiguity about all of the states simultaneously

A quantum system as superposition state in indefinite Hilbert space

3. Cognitive measures disturb each other, creating uncertainty

Views about climate change vs. employment

- If you ask someone directly about climate change, they may say its important and needs to be addressed
- However, if you *first* ask how important it is to keep low levels of unemployment, they may become less certain about the need to address climate change

- → Questions are incompatibile – we can’t answer questions simultaneously, and one question disturbs the answer to another, so that judgments do not commute

- Quantum measurement disturbs/changes the system
- Quantum operators do not commute in general

4. Cognitive logic does not obey classic logic

Classic Information Processing

- You may believe a person is guilty or not guilty (two mutually exclusive and exhaustive events)
- You may feel a person is good or bad (two mutually exclusive and exhaustive events)
- Distributive Axiom of Boolean Logic applies

$$\text{Guilty} \wedge (\text{Good} \vee \text{Bad})$$
$$= (\text{Guilty} \wedge \text{Good}) \vee (\text{Guilty} \wedge \text{Bad})$$

Quantum Information Processing

- Existence of a superposed state

→ Distributive Axiom does not always apply

$$\text{Guilty} \wedge (\text{Good} \vee \text{Bad})$$
$$\neq (\text{Guilty} \wedge \text{Good}) \vee (\text{Guilty} \wedge \text{Bad})$$

→ Human judgments do not necessary obey classic Boolean logic, and comply with a more general quantum logic.

5. Classical cognitive models cannot account for apparent non-compositionality of concept combinations

apple chip

How do we ascribe meaning to such novel conceptual combinations?

How..? Semantic compositionality

The Principle of Semantic Compositionality (sometimes called 'Frege's Principle')

Whole = sum of the parts

…. Are conceptual combinations such as "apple chip" semantically compositional?

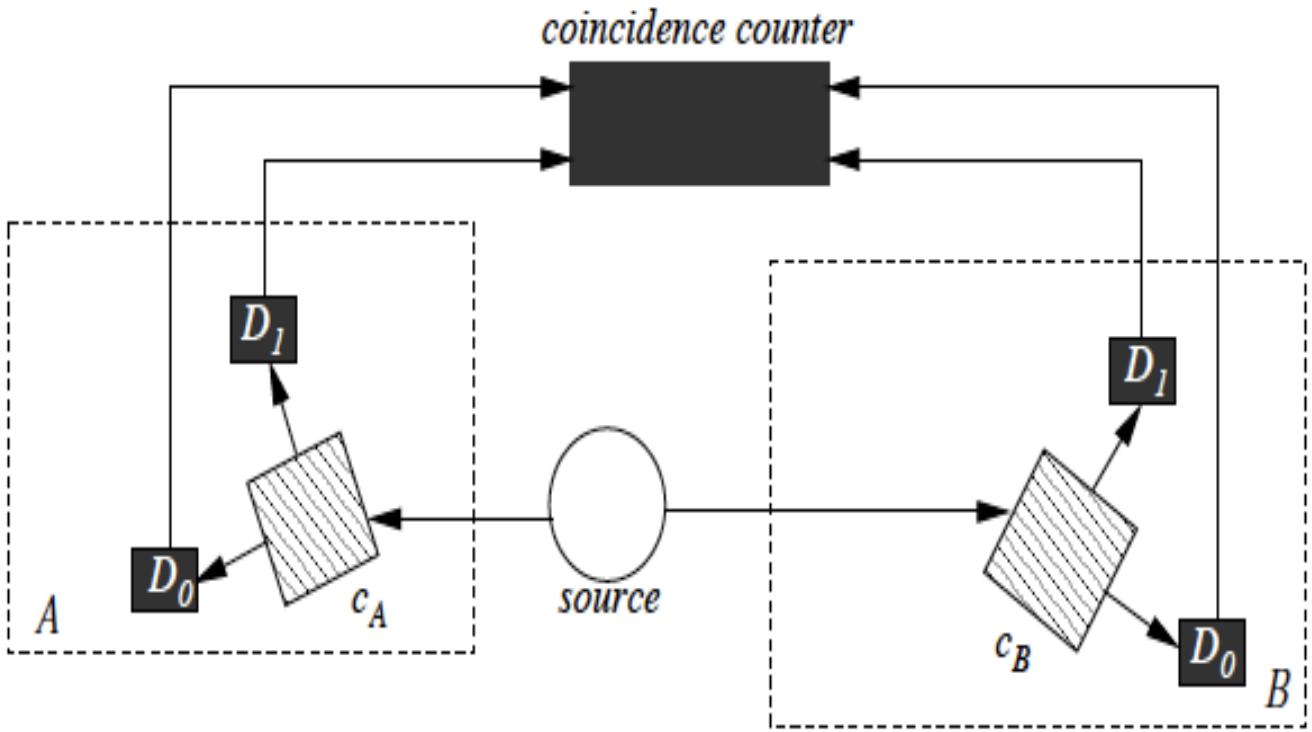
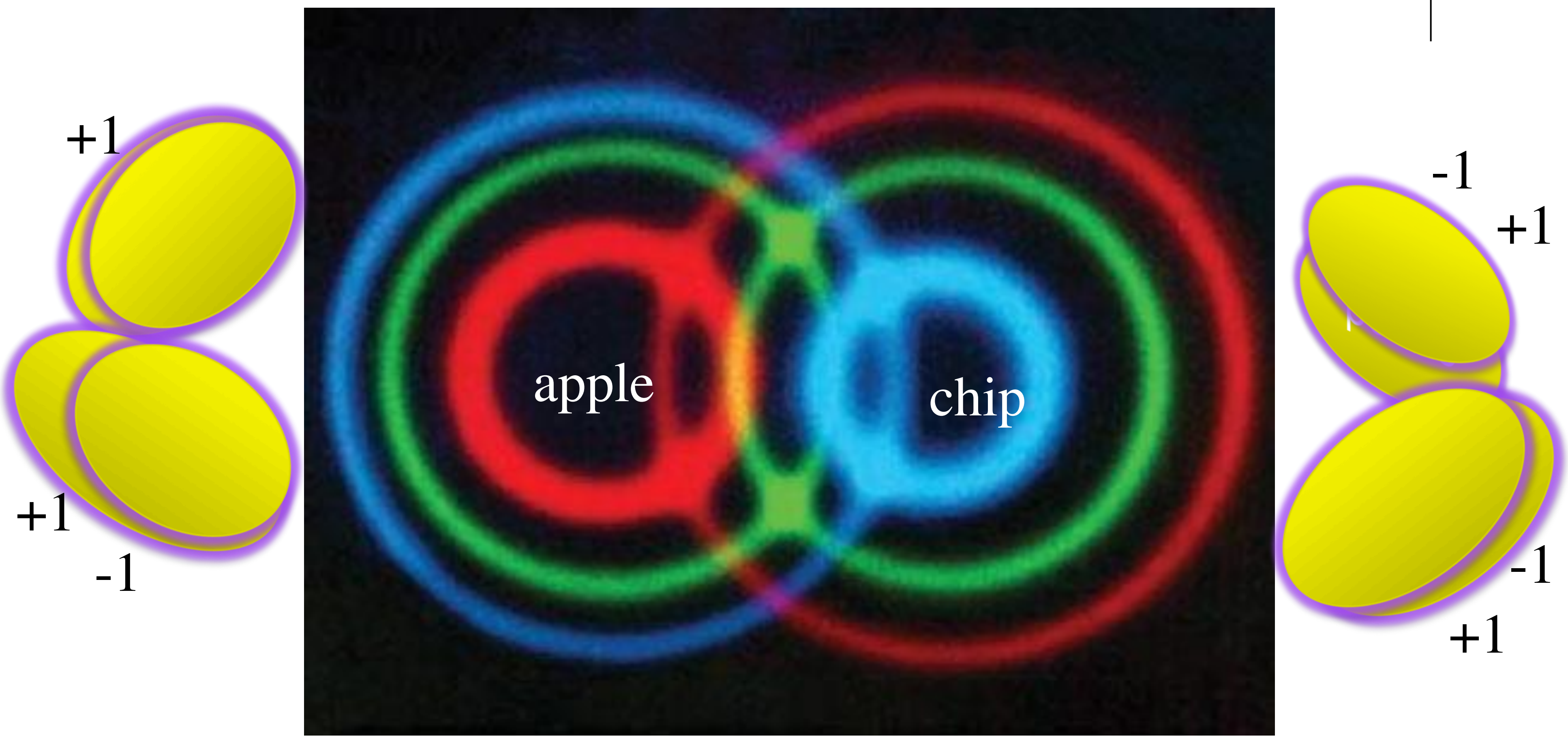


Figure 4: An experimental scenario testing for the non-decomposability of an entangled system of two polarized photons. A source emits two entangled photons that travel to polarizers at c_A and c_B. In each of the regions A and B, either detector D_0 (polarization is "down") or D_1 (polarization is "up") clicks, and this is recorded at a coincidence counter.

Conceptual combination ≈ entangled photons:

A, B concepts, polarization ≈ sense (e.g. fruit sense of "apple")
Polarizers ≈ primes to orient interpretation (e.g., "banana" orients subject to fruit sense of "apple")

Bridge: conceptual combination ≈ entangled photons



“a nano-chipped granny smith”
“dried pieces of apple that you eat”

Definition of Entanglement

Definition 1. (QE): Let A be an n-qubit system in a state $|\phi_A\rangle$ and $\{A_1, A_2\}$ be a partition of A, where two disjoint parts A_1 and A_2 have $0 < k < n$ qubits and $n - k$ qubits, respectively. A is entangled iff. there does NOT exist any tensor product decomposition of $|\phi_A\rangle$ such that $|\phi_A\rangle = |\phi_{A_1}\rangle \otimes |\phi_{A_2}\rangle$, where $|\phi_{A_1}\rangle$ and $|\phi_{A_2}\rangle$ are the states of A_1 and A_2, respectively.

- **a system can not be factorized into two subsystems via tensor product.**

Detecting entanglement

$$
\begin{array}{cc|cc}
 & & \multicolumn{2}{c}{B} \\
 & & \mathbf{B1} & \mathbf{B2} \\
 & & \begin{array}{cc} +1 & -1 \end{array} & \begin{array}{cc} +1 & -1 \end{array}
\end{array}
$$

$$
A\quad
\begin{array}{cc}
\mathbf{A1} & \begin{array}{c} +1 \\ -1 \end{array} \\
\mathbf{A2} & \begin{array}{c} +1 \\ -1 \end{array}
\end{array}
\left(\begin{array}{cc|cc}
p_1 & p_2 & p_5 & p_6 \\
p_3 & p_4 & p_7 & p_8 \\
\hline
p_9 & p_{10} & p_{13} & p_{14} \\
p_{11} & p_{12} & p_{15} & p_{16}
\end{array}\right)
$$

Can we construct a joint probability distribution Pr(**A1,A2,B1,B2**) from from the pair-wise joint distributions which are empirically collected:
Pr(**A1,B1**),Pr(**A1,B2**),Pr(**A2,B1**), Pr(**A2,B2**)

CHSH inequality:

$$|E(A1\ B1) + E(A1\ B2) + E(A2\ B1) - E(A2\ B2)| \le 2$$

Detecting entanglement

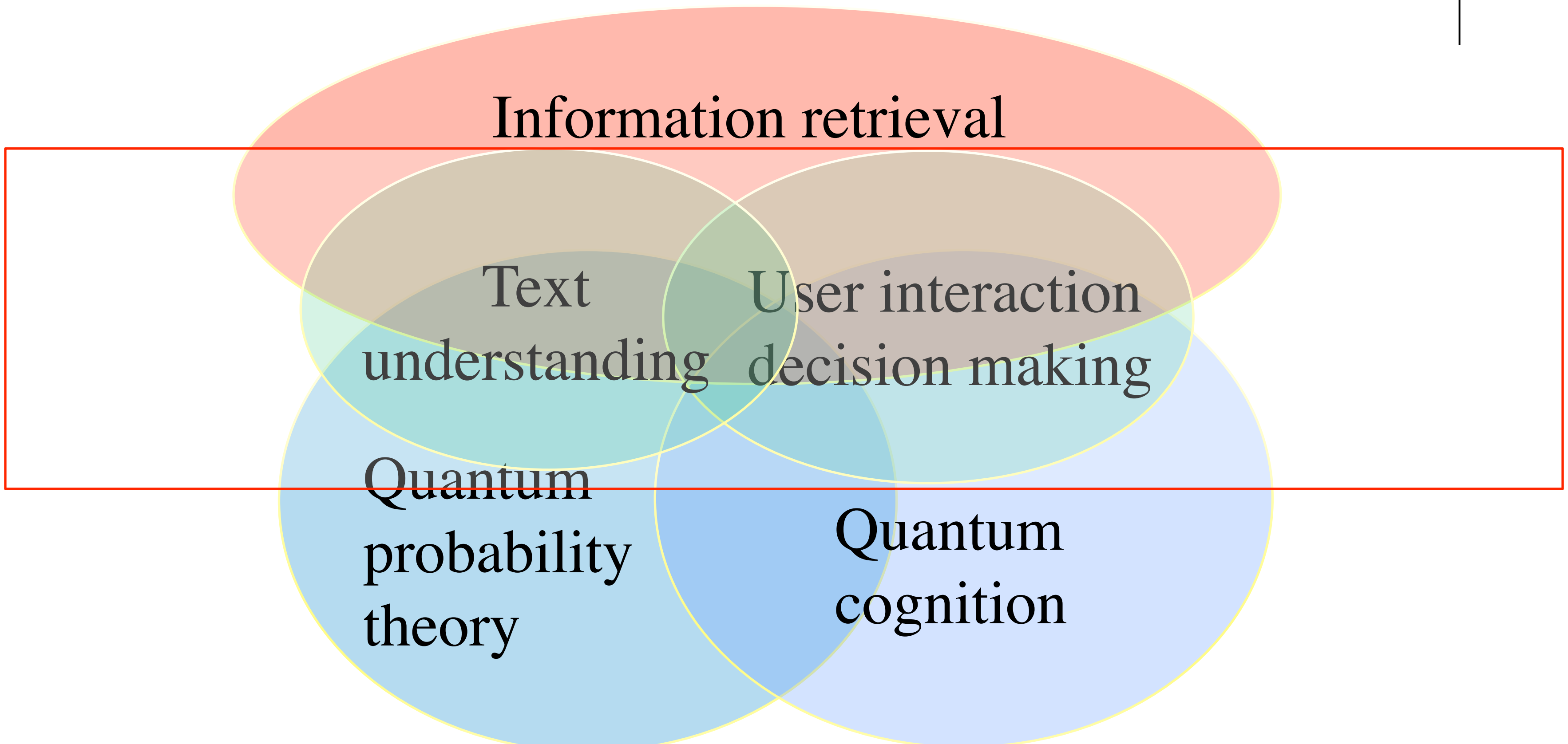
			CHIP			
			B1(*potato*)		**B2**(*circuit*)	
			+1	−1	+1	−1
APPLE	**A1**(*banana*)	+1	0.94	0.06	0	0.75
		−1	0	0	0.25	0
	A2(*computer*)	+1	0	0.35	0.47	0
		−1	0.65	0	0	0.53

N=65

CHSH > 2
(non-compositional)

Quantum probabilistic model required

A Quantum Cognitive View of IR



- "quantumeness" of users in relevance decision making
- Text representation in line with human understanding of text

Case Study 1

Quantum-like Contextuality in Relevance decision making

Relevance

- A cognitive concept
- A relation. Relevance “to” something
- Relates an information object to a context or a situation (information need)
- Fundamentally contextual
  - No pre-defined and fixed values of relevance
  - Explicit, causal: e.g. “things near me”- time, location, weather
  - Implicit: e.g. relevance judgment of other documents

Quantum Contextuality

- Contextuality is a fundamental feature of quantum systems
- Quantum contextuality exists due to the inherently random nature of systems, rather than direct, causal influences.
- Impossible to pre-assign a value to the property of a system independent of the context.
- The value of a property comes into existence only at the instance of measurement.
- Contradiction with classical world:
  - Any measurable property (for example, weight of a person) has values independent of measurement and measurement only serves to reveal its value.
  - Any measure of randomness is only due to ignorance of certain latent variables of the system.
  - Quantum Mechanics was proven incompatible with hidden variable theories.
- Impossibility of assigning joint probability distribution P(A1,A2,B1,B2) with the marginal distributions P(A1, A2) and P(B1, B2) obtained from different measurement contexts.

Definition of Entanglement

Definition 1. (QE): Let A be an n-qubit system in a state $|\phi_A\rangle$ and $\{A_1, A_2\}$ be a partition of A, where two disjoint parts A_1 and A_2 have $0 < k < n$ qubits and $n - k$ qubits, respectively. A is entangled iff. there does NOT exist any tensor product decomposition of $|\phi_A\rangle$ such that $|\phi_A\rangle = |\phi_{A_1}\rangle \otimes |\phi_{A_2}\rangle$, where $|\phi_{A_1}\rangle$ and $|\phi_{A_2}\rangle$ are the states of A_1 and A_2, respectively.

- **a system can not be factorized into two subsystems via tensor product.**

Detecting entanglement

$$
\begin{array}{ccc}
 & & B \\
 & & \begin{array}{cccc} \multicolumn{2}{c}{\mathbf{B1}} & \multicolumn{2}{c}{\mathbf{B2}} \\ +1 & -1 & +1 & -1 \end{array} \\
A & \begin{array}{cc} \mathbf{A1} & \begin{array}{c} +1 \\ -1 \end{array} \\ \mathbf{A2} & \begin{array}{c} +1 \\ -1 \end{array} \end{array} & \left(\begin{array}{cc|cc} p_1 & p_2 & p_5 & p_6 \\ p_3 & p_4 & p_7 & p_8 \\ \hline p_9 & p_{10} & p_{13} & p_{14} \\ p_{11} & p_{12} & p_{15} & p_{16} \end{array}\right)
\end{array}
$$

Can we construct a joint probability distribution Pr(**A1,A2,B1,B2**) from the pair-wise joint distributions which are empirically collected:
Pr(**A1,B1**),Pr(**A1,B2**),Pr(**A2,B1**), Pr(**A2,B2**)

CHSH inequality:
$|E(A1\ B1) + E(A1\ B2) + E(A2\ B1) - E(A2\ B2)| \leq 2$

Contextuality by Default Theory

- Entanglement considered as contextuality at a distance
- CHSH inequality in quantum mechanics assumes no signalling
- In human cognition, there may be signalling (direct influence), denoted as Δ

$|E(A1B1) + E(A2B1) + E(A2B1) - E(A2B2)| - \Delta \leq 2$

$\Delta = |E[]-E[]|+|E[]-E[]|+|E[]-E[]|+|E[]-E[]|$

$$s_{odd}(E[R_1^2 R_2^1], E[R_2^3 R_3^2], \ldots, E[R_n^1 R_1^n]) - (n-2) - \Delta > 0$$

$$s_{odd}(x_1, \ldots, x_n) = \max(\pm x_1 \pm \ldots \pm x_n)$$

Ehtibar N. Dzhafarov, et al. 2017. Contextuality in canonical systems of random variables. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 375, 2106 (oct 2017).
Ehtibar N. Dzhafarov, et al. 2016. Contextuality-by-Default: A Brief Overview of Ideas, Concepts, and Terminology. In Quantum Interaction, 12–23.

Experiment - Material

- Three queries from TREC 2013 Webtrack
- Three document snippets were selected for each query.
- Documents were paired to form three contexts. For example, for documents D1, D2 and D3, we created three contexts - {D1, D2}, {D2, D3} and {D3, D1}.
- Users were asked to decide which document is more relevant to the query
- 242 participants using crowdsourcing platform – Prolific.ac

Experiment - Query

Query 1	hawaiian volcano observatories
Description:	I am looking for the history of and a summary of the work performed at the Hawaiian Volcano Observatories.
Query 2	hurricane Irene flooding in manville nj
Description	How has the flooding that resulted from hurricane Irene affected Manville, NJ?
Query 3	frank lloyd wright biography
Description	Find biographical information for Frank Lloyd Wright.

Experiment – Interface

Description: I am looking for the history of and a summary of the work performed at the Hawaiian Volcano Observatories.

Search Query: hawaiian volcano observatories

Document 1:

Hawaiian Volcano Observatory Adapts to Recent Changes | Big Island ...
bigislandnow.com/2018/11/.../hawaiian-volcano-observatory-adapts-to-recent-change... ▾
Nov 16, 2018 - The USGS **Hawaiian Volcano Observatory** continues to closely monitor volcanoes and earthquakes on the Island of Hawai'i. On this map ...

Document 2:

Hawaiian Volcano Observatory - USGS: Volcano Hazards Program
https://volcanoes.usgs.gov/observatories/hvo/ ▾
Through VNS, the Hawaiian Volcano Observatory issues: daily **Kīlauea** eruption updates, weekly **Mauna Loa** updates, monthly updates for **Hualālai**, **Haleakalā**, and **Mauna** Kea, Status Reports about volcanic activity during ongoing events,

Question:

*Which of the two documents you think is more **relevant** to the **search query**?*

Choose one appropriate option from below:

- ◯ Document 1
- ◯ Document 2

Experiment - Results

		D_2	
		m	l
D_1	m	0	0.775
	l	0.225	0

Context 1: Order D_1, D_2

		D_1	
		m	l
D_2	m	0	0.2927
	l	0.7073	0

Context 1: Order D_2, D_1

		D_3	
		m	l
D_2	m	0	0.2927
	l	0.7073	0

Context 2: Order D_2, D_3

		D_2	
		m	l
D_3	m	0	0.625
	l	0.375	0

Context 2: Order D_3, D_2

		D_1	
		m	l
D_3	m	0	0.462
	l	0.538	0

Context 3: Order D_3, D_1

		D_3	
		m	l
D_1	m	0	0.525
	l	0.475	0

Context 3: Order D_1, D_3

- No significant order effect between documents in the pair -> existence of marginal joint probability distributions

- CbD inequality violated -> Existence of implicit contextuality

Query	Δ
Query 1	0.9629
Query 2	1.400
Query 3	1.7999

Case Study 2:

Complex Hilbert Space of Multi-dimensional Relevance

Multidimensional Relevance

- Manifestation in terms of judgement criteria (Dimensions)
  - Traditionally considered to be Topical
  - Other factors or dimensions affecting relevance.
  - Reliability/Credibility, Understandability/Readability, Novelty/Diversity, Interest, etc.

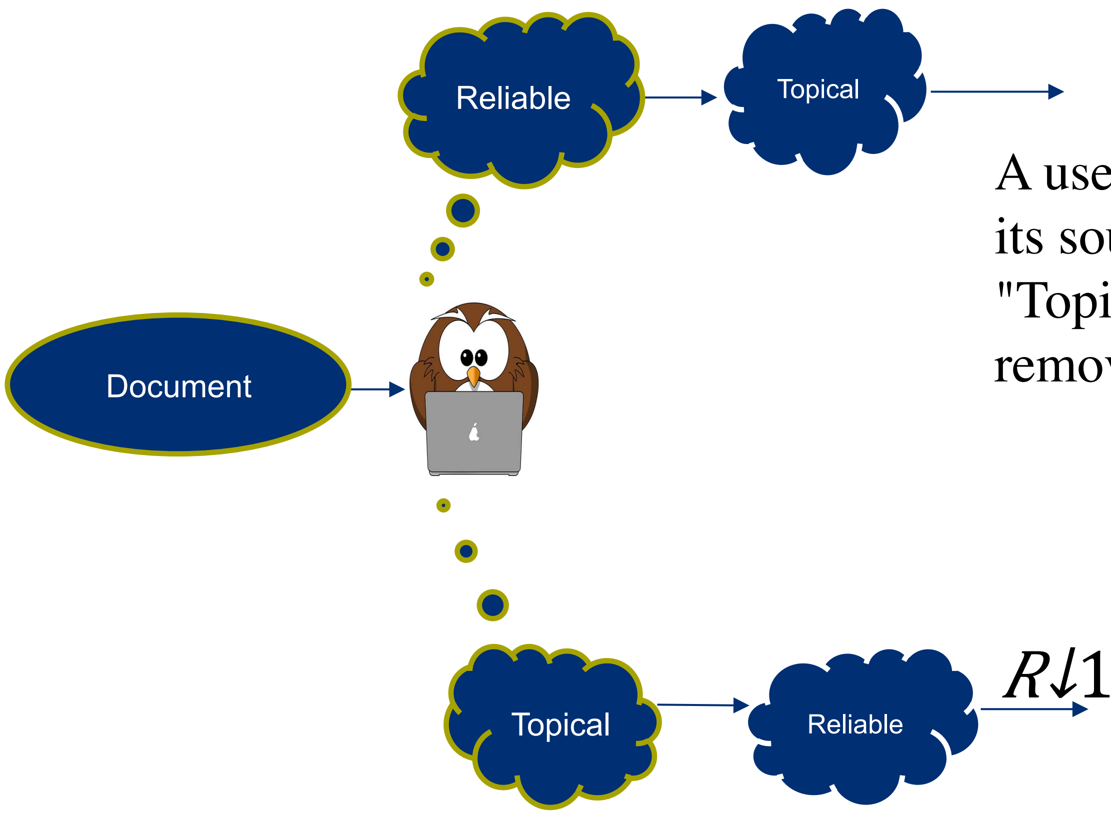
- Different manifestations are considered as layers interacting with each other[1].
- Each of these interacting layers include inferences of relevance.

[1] Tefko Saracevic. 1997. The stratified model of information retrieval interaction :Extension and applications.

Multidimensional Relevance

- In a search session or search task, there is combination (and ordering) of dimensions which user has in mind for judging documents.
- For query “Visa to the USA”, “Topicality” and “Reliability” maybe the predominant factors to judge documents
- For “Hotels in Palo Alto”, the user might go to the preferred websites that give a wide range of selections (“Habit” and “diversity” dimensions more important)
- For some randomly sampled **4837** sessions of Bing query log, we found that in **3910** or **80.84%** of the sessions, one of the top three dimensions for the first query of the session remains in the top three for all the queries of the session.

Incompatible Decision Perspectives



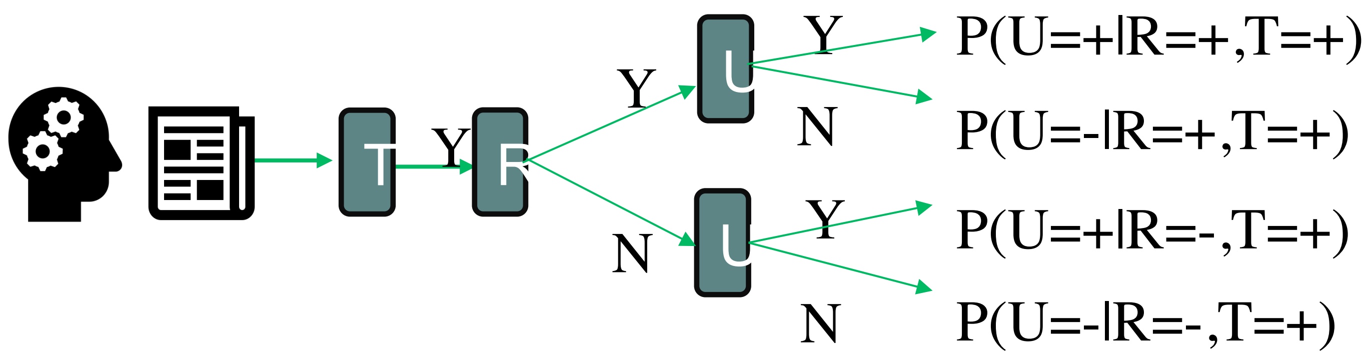
A user may find a document less reliable due to its source, but when the user considers the "Topicality" dimension and reads it, it might remove the doubts about the reliability.

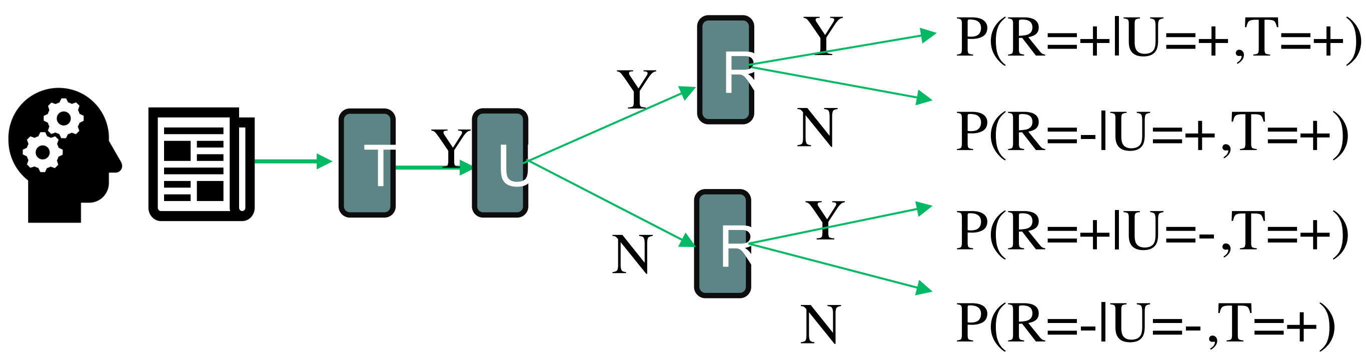
In the case a user considers multiple relevance dimensions for judging a document, the final judgment will depend upon the order of dimensions considered, called Order Effect in Psychology.

Research Questions

1) How does consideration of one relevance dimension affect inference of relevance with respect to another dimension.
2) Can we construct a formal mathematical model of the user's underlying cognitive state in order to make predictions about such interactions?

Experiment - Design





Experiment - Material

- 3 query-document pairs.
- Between-subject design: 2 groups of participants
- 300 participants using crowdsourcing platform – Prolific.ac
- Documents modified so as to introduce uncertainty in the judgement of these dimensions.

Experiment

- Experiment to study interaction between relevance dimensions.
- We study 3 dimensions - Topicality, Understandability, Reliability
- Asking user yes/no questions about dimensional relevance in different orders.

Experiment - Query

	Query 1	Query 2	Query 3
Query Terms	radio waves and brain cancer	symptoms of mad cow disease in humans	educational advantages of social networking sites
Information Need	Look for evidence that radio waves from radio towers or mobile phones affect brain cancer occurrence.	Find information about mad cow disease symptoms in humans.	What are the educational benefits of social networking sites?
Source	TREC 2005 Robust Track (310)	TREC 2013 Web Track (236)	TREC 2014 Web Track (293)

Experiment – Document design

Information Need: Look for evidence that radio waves from radio towers or mobile phones affect brain cancer occurrence.

Query: radio waves and brain cancer

Here are some possible radiation dangers in your environment....

https://hustlebustlenews.com/here-are-some-radiation..

May 5, 2014 - A study examined the role of occupational RF/ MW-EMF exposure in the risk of meningioma.....the International Commission on Non-Ionizing Radiation Protection. Several conditional logistic regressions performed for glioma and meningioma. No significant association.....However, the slight increase in risk...merits further research....

Complex Hilbert Space of Cognitive State

Representing User Cognitive State:

$$|S\rangle = t|T+\rangle + \sqrt{1-t^2}\,|T-\rangle$$

$|T+\rangle$ - State of judging document as topically relevant

$|T-\rangle$ - State of judging document as topically non-relevant

$|S\rangle$ - User's state before the judgement

t^2 - Probability of judging document as topically relevant

Complex-valued Hilbert Space

Representing User Cognitive State:

$$|U+\rangle = u|T+\rangle + \sqrt{1-u^{2}}\,|T-\rangle$$

$$|R+\rangle = r|T+\rangle + \sqrt{1-r^{2}}\,e^{i\theta_{r}}\,|T-\rangle$$

	Query 1	Query 2	Query 3
	0.7622	0.6736	0.8993
	0.5779	0.8041	0.9701
	0.5462	0.7311	0.6456
	80.62 deg	56.79 deg	51.43 deg

$$P_{U+T+} = |\langle U+|T+\rangle|^{2} = u^{2}$$

$$P_{R+T+} = |\langle R+|T+\rangle|^{2} = r^{2}$$

$$P_{R+U+,\,T+} = |\langle R+|U+\rangle|^{2} = f(u, r, \theta_{r})$$

$$= (ur)^{2} + (1-u^{2})(1-r^{2}) + 2ur\sqrt{(1-u^{2})(1-r^{2})}\cos\theta_{r}$$

Experiment - Analysis

Effect of Understandability on Reliability

	Q1	Q2	Q3
	0.5462	0.7311	0.6456
	0.5872	0.8332	0.7384
	0.3692	0.5261	**0.0000**

Effect of Reliability on Understandability

	Q1	Q2	Q3
	0.5779	0.8040	0.9701
	0.5999	0.8822	0.9633
	0.4074	**0.4801**	0.8887

Law of total probability (LTP) violated:

P(R+,T+) = P(R+, U+, T+) + P(R+, U-, T+)

$$P_u(R+, T+) = P(R+, U+, T+) + P(R+, U-, T+) + Int(\theta_r)$$

	P(R+,T+)	P_u(R+,T+)
Query 1	0.3775	0.4609
Query 2	0.5207	0.4857
Query 3	0.6442	0.5616

Experiment - Analysis

•

- Non-commutativity of operators implies presence of order effects.
- **$P(T,R) \neq P(R,T)$**
- Even though we do not ask questions in T,R and R,T order, we can predict an order effect between these two dimensions!

Reflections

- A Complex Hilbert Space representation of user's cognitive state for a document is introduced, inspired from Quantum Physics.
- In general there is incompatibility in relevance dimensions
- A Quantum probabilistic explanation of the Order effects arising from incompatibility is provided.

Case Study 3:

Complex Hilbert Space of Text Representation Learning

Semantic Hilbert Space for text representation learning

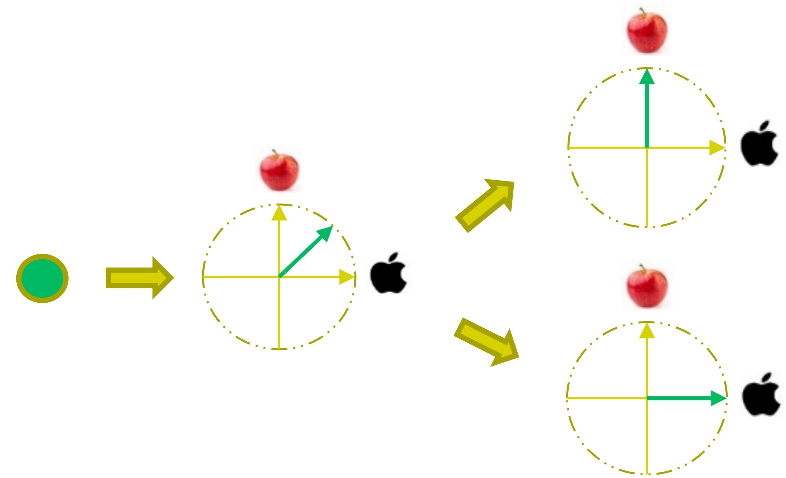
- A unified quantum view of different levels of linguistic units
  - **Sememes → basis states** $\{|e_j\rangle\}_{j=1}^{n}$ **. Basis of Semantic Hilbert Space**
  - **Words → superposition states** $|w_l\rangle = \sum_{j=1}^{n} r_{l,j}\, e^{i\phi_{l,j}} |e_j\rangle$
  - **Word Composition → density matrix** $\rho = \sum_{w_i \in d} p(i) |w_i\rangle\langle w_i|$
  - **High Level Features → measurement probabilities** $\{tr(\rho |v_j\rangle\langle v_j|)\}_{j=1}^{k}$

Uncertainty in Language/QT

- A **single word** may have multiple meanings

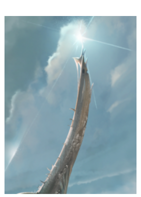
"Apple"

- Uncertainty of a pure state



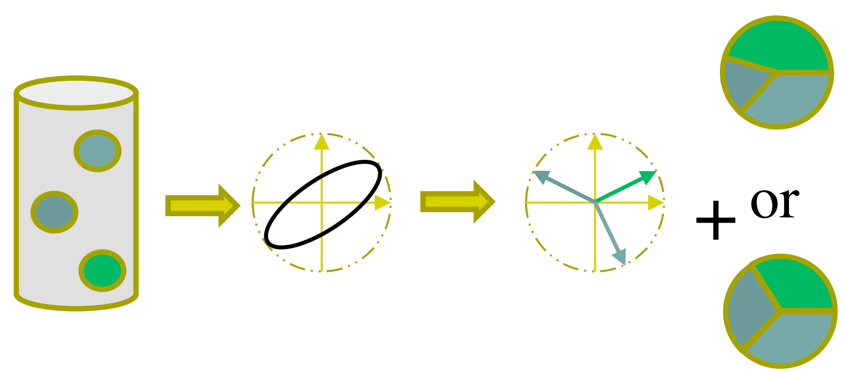
- **Multiple words** may be combined in different ways

"Ivory tower"

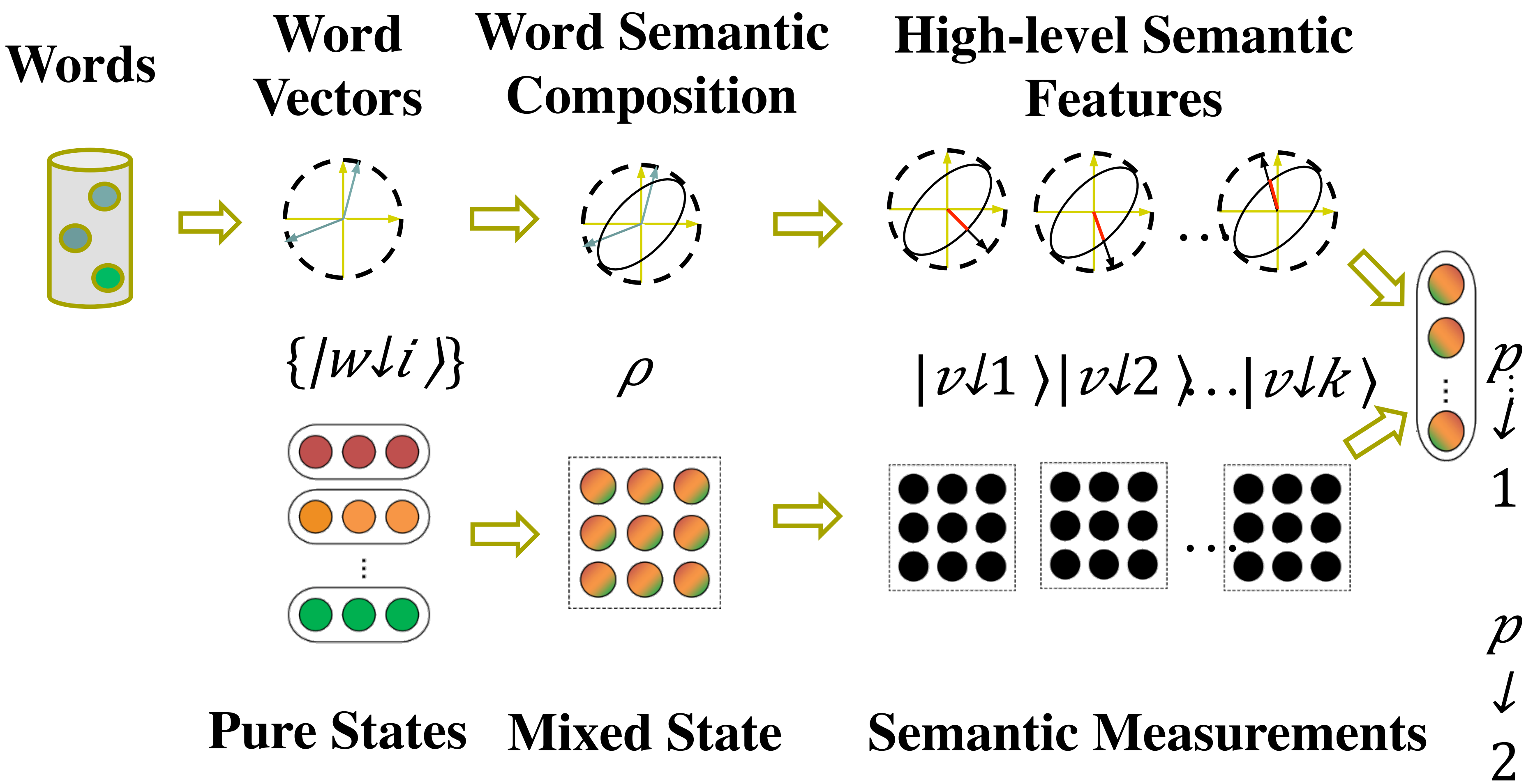


UNIVERSITY OF MINNESOTA

- Uncertainty of a mixed state



Semantic Hilbert Space



Benyou Wang*, Qiuchi Li*, Massimo Melucci, and Dawei Song. *Semantic Hilbert Space for Text Representation Learning*. In *WWW2019*

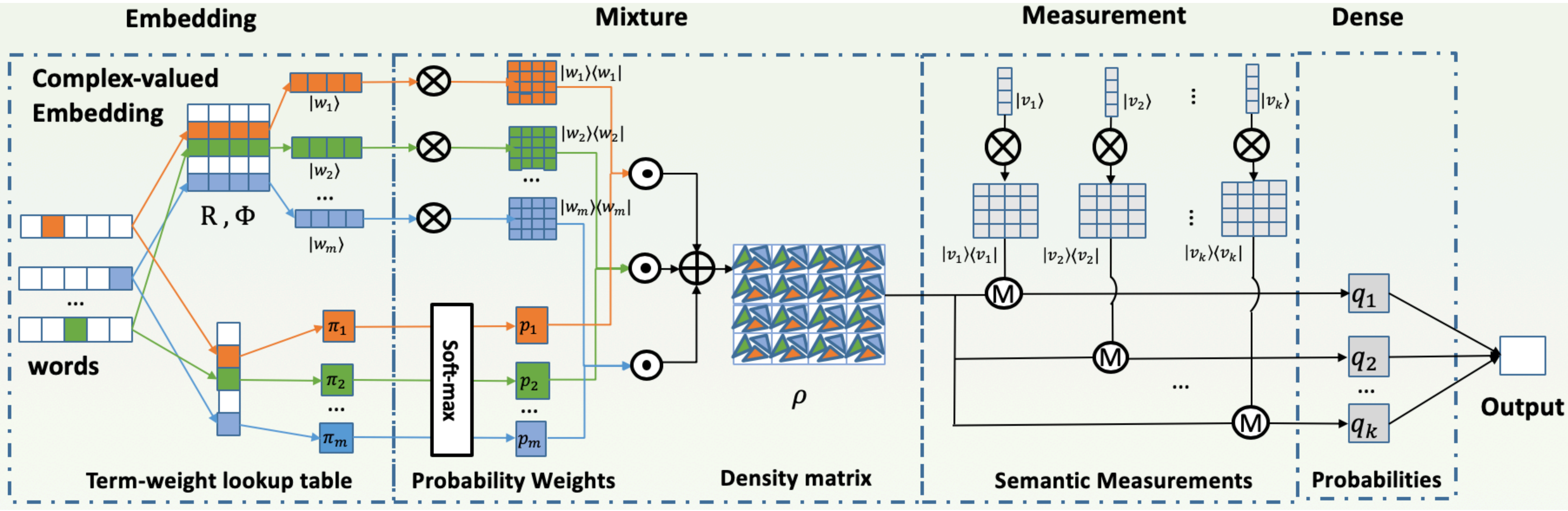
Semantic Hilbert Space

- Complex-valued word embedding gives rise to non-linear combination of word features
    - $W_l = [r_{l,1} e^{i\phi_{l,1}}, \ldots, r_{l,n} e^{i\phi_{l,n}}]$, amplitudes $r_{l,1}, \ldots, r_{l,n}$ and phases $\phi_{l,1}, \ldots, \phi_{l,n}$ carries different levels of information

    - The mixture process implicitly performs a non-linear combination of amplitudes and phases *within a single word* as well as *between different words*
        - $re^{i\phi} = r_1 e^{i\phi_1} + r_2 e^{i\phi_2} \rightarrow \begin{cases} r = \sqrt{|r_1|^2 + |r_2|^2 + 2 r_1 r_2 \cos(\phi_1 - \phi_2)} \\ \phi = \arctan\left(\frac{r_1 \sin\phi_1 + r_2 \sin\phi_2}{r_1 \cos\phi_1 + r_2 \cos\phi_2}\right) \end{cases}$
- The semantic measurements are trainable, enabling one to find discriminative measurement projectors in a data-driven way

Semantic Hilbert Space

Quantum Probability Driven Network for Text Classification



Semantic Hilbert Space

Results

Model	CR	MPQA	MR	SST	SUBJ	TREC
Uni-TFIDF	79.2	82.4	73.7	-	90.3	85.0
Word2vec	79.8	**88.3**	77.7	79.7	90.9	83.6
FastText [18]	78.9	87.4	76.5	78.8	91.6	81.8
Sent2Vec [24]	79.1	87.2	76.3	80.2	91.2	85.8
CaptionRep [15]	69.3	70.8	61.9	-	77.4	72.2
DictRep [16]	78.7	87.2	76.7	-	90.7	81.0
Ours: QPDN	**81.0**†	87.0	**80.1**†	**83.9**†	**92.7**†	**88.2**†
CNN [19]	81.5	89.4	81.1	88.1	93.6	92.4
BiLSTM [10]	81.3	88.7	77.5	80.7	89.6	85.2

Semantic Hilbert Space

Explainability

Components	DNN	QPDN
Sememe	-	basis vector / **basis state** $\{w \mid w \in C^n, \|\|w\|\|_2 = 1, \}$ complete &orthogonal
Word	real vector $(-\infty, \infty)$	unit complex vector / **superposition state** $\{w \mid w \in C^n, \|\|w\|\|_2 = 1\}$
Low-level representation	real vector $(-\infty, \infty)$	density matrix / **mixed system** $\{\rho \mid \rho = \rho^*, tr(\rho) = 1\}$
Abstraction	CNN/RNN $(-\infty, \infty)$	unit complex vector / **measurement** $\{w \mid w \in C^n, \|\|w\|\|_2 = 1\}$
High-level representation	real vector $(-\infty, \infty)$	probabilities/ **measured probability** $(0, 1)$

- With well-constraint complex values, our model components can be **explained** as concrete **quantum states** at design phase

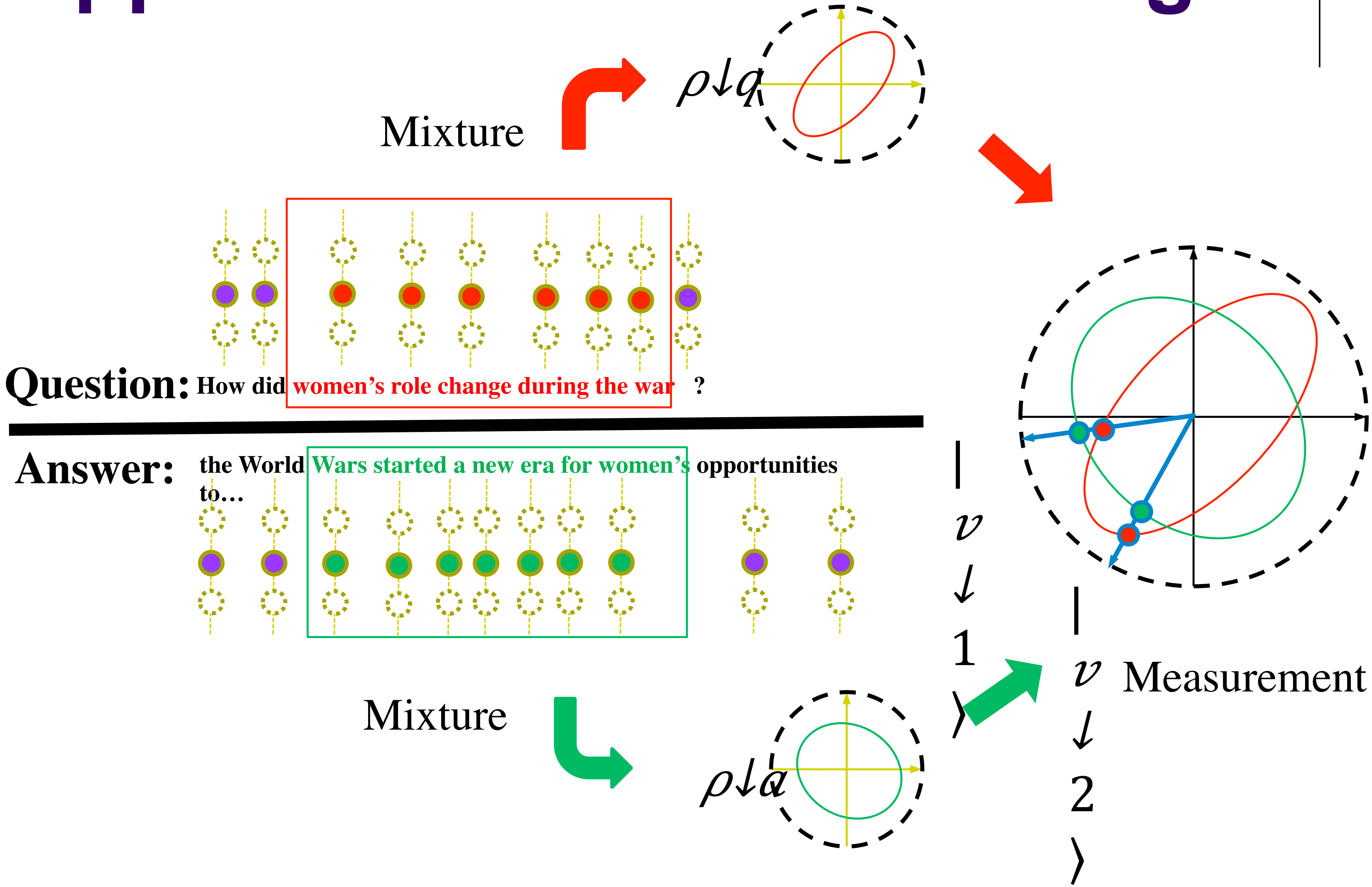
Semantic Hilbert Space

Trainable semantic measurements

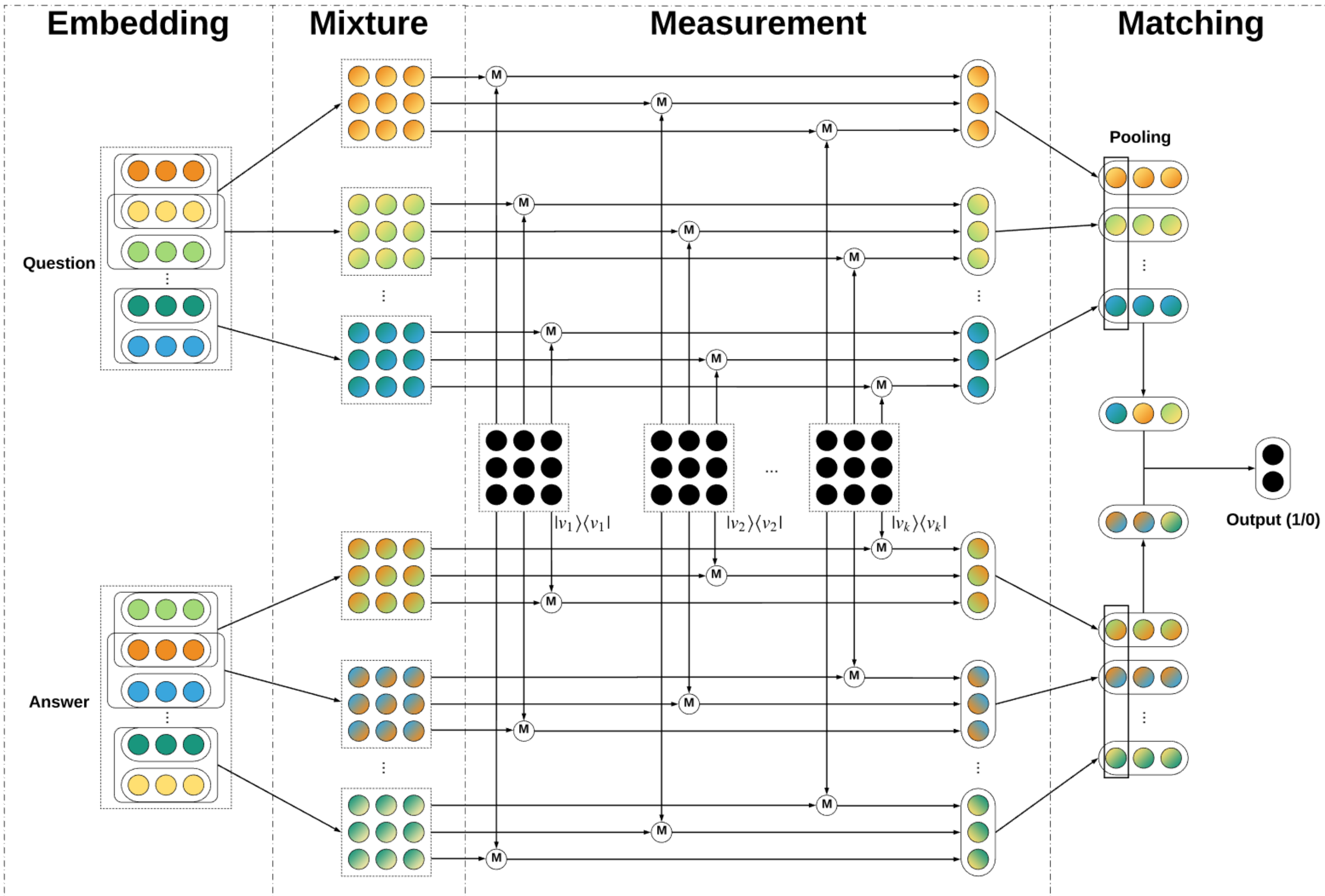
- Semantic measurements are superposition states in the *same Semantic Hilbert Space*
- We can understand measurements by referring to their neighboring words, while it is not easy for CNN/RNN cells

Measurement	Selected neighborhood words
1	change, months, upscale, recently, aftermath
2	compelled, promised, conspire, convince, trusting
3	goo, vez, errol, esperanza, ana
4	ice, heal, blessedly, sustains, make
5	continue, warned, preposterousness, adding, falseness

Application to Text Matching



Complex-valued Network for Matching



Experiment Result

- Effectiveness
  - Competitive compared to strong baselines
  - Outperforms existing quantum-inspired QA model

Model	MAP	MRR
Bigram-CNN	0.5476	0.6437
LSTM-3L-BM25	0.7134	0.7913
LSTM-CNN-attn	0.7279	0.8322
aNMM	0.7495	0.8109
MP-CNN	**0.7770**	0.8360
CNTN	0.7278	0.7831
PWIM	0.7588	0.8219
QLM	0.6780	0.7260
NNQLM-I	0.6791	0.7529
NNQLM-II	0.7589	0.8254
CNM	0.7701	**0.8591**
Over NNQLM-II	1.48% ↑	4.08% ↑

Experiment Results on TREC QA Dataset. The best performing values are in bold.

Model	MAP	MRR
Bigram-CNN	0.6190	0.6281
QA-BILSTM	0.6557	0.6695
AP-BILSTM	0.6705	0.6842
LSTM-attn	0.6639	0.6828
CNN-Cnt	0.6520	0.6652
QLM	0.5120	0.5150
NNQLM-I	0.5462	0.5574
NNQLM-II	0.6496	0.6594
CNM	**0.6748**	**0.6864**
Over NNQLM-II	3.88% ↑	4.09% ↑

Experiment Results on WikiQA Dataset.The best performing values are in bold.

Semantic Hilbert Space

- Reflections
  - Interpretability for language understanding
  - Quantum-inspired complex-valued network
    - Transparent & Post-hoc Explainable
    - Comparable to strong baselines

Case Study 4:

An End-to-End Quantum-like language Model

Neural Network based Quantum-like Language Model(NNQLM)

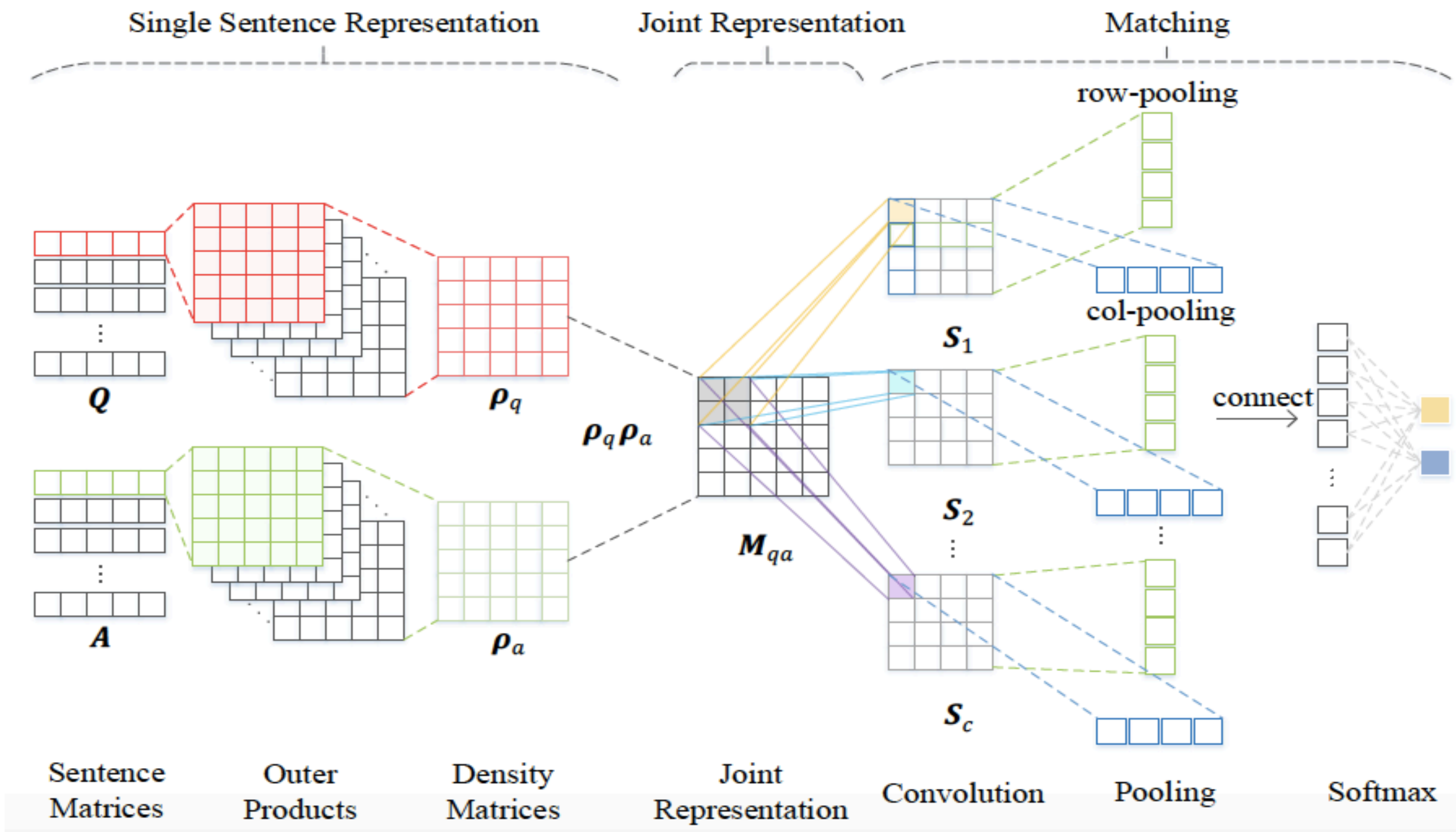
Quantum Language Model (QLM)

- A sequence of quantum events.
- The probability uncertainty of single terms or term dependencies, are encoded as a **density matrix** ρ

Our Approach

- Embedding vector as the Input for the global semantics
- Analytical solution for density matrix estimation
- End-to-End QLM based on Convolutional Neural Network

Neural Network based Quantum-like Language Model(NNQLM)



Single Sentence Representation (Density Matrix with Embedding as States)

$$\rho = \sum_i p_i S_i = \sum_i p_i |s_i\rangle\langle s_i|$$

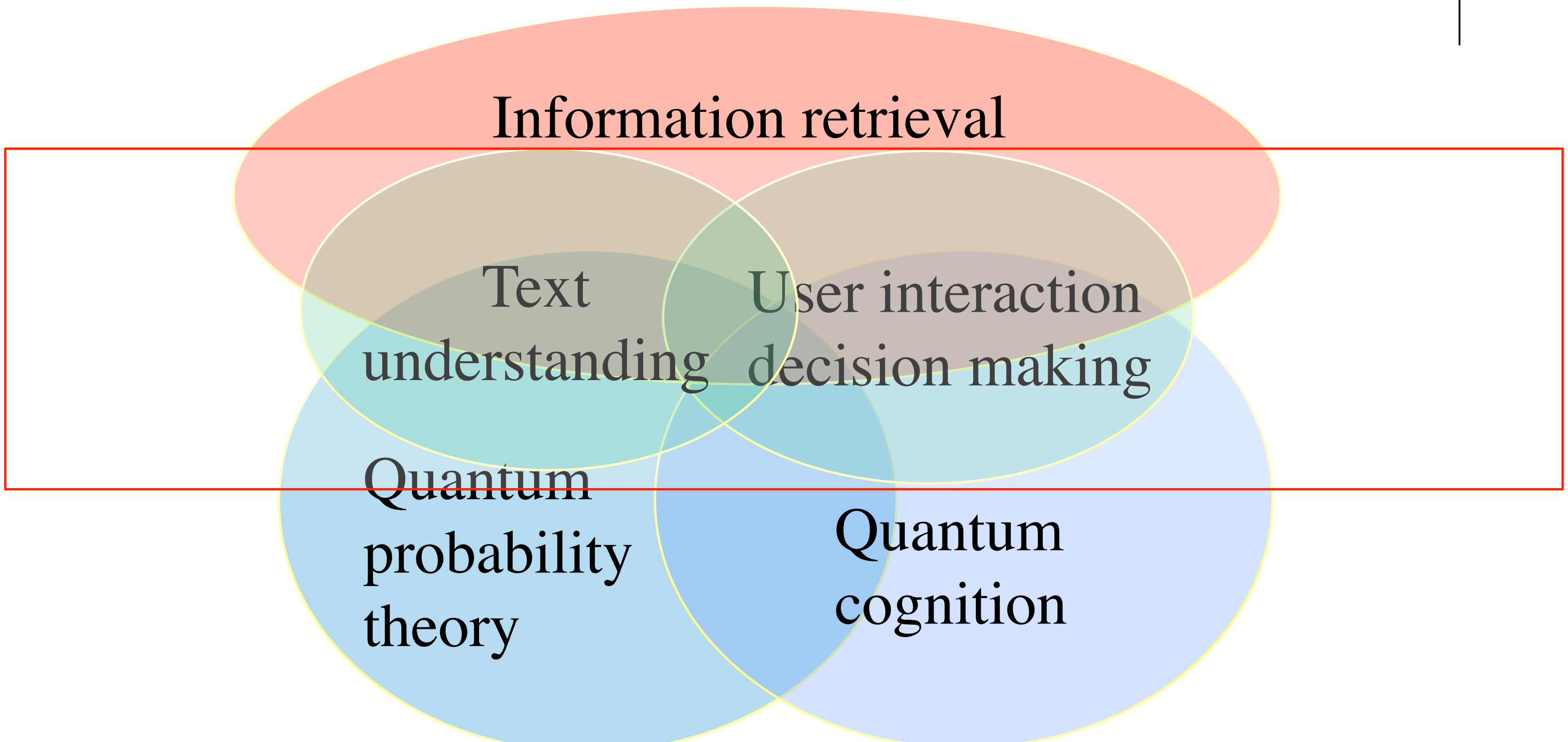
Joint Representation (Encoding the Similarity Features Between Q and A)

$$M_{qa} = \rho_q \rho_a = \sum_{i,j} \lambda_i \lambda_j |r_i\rangle\langle r_i | r_j\rangle\langle r_j|$$

Convolution Neural Network (Extracting Richer Similarity Patterns)

NNQLM achieves significant improvements over the original QLM and a comparable result to the state-of-the-art approaches on Question Answering

A Quantum Cognitive View of IR



- "quantum-like structures" exist in user's cognitive state in relevance decision making and text understanding with uncertainty
- Quantum probability theory, which is a generalization of classical theory, is necessary to cope with the "quantumness" of users

More Quantum-inspired IR models

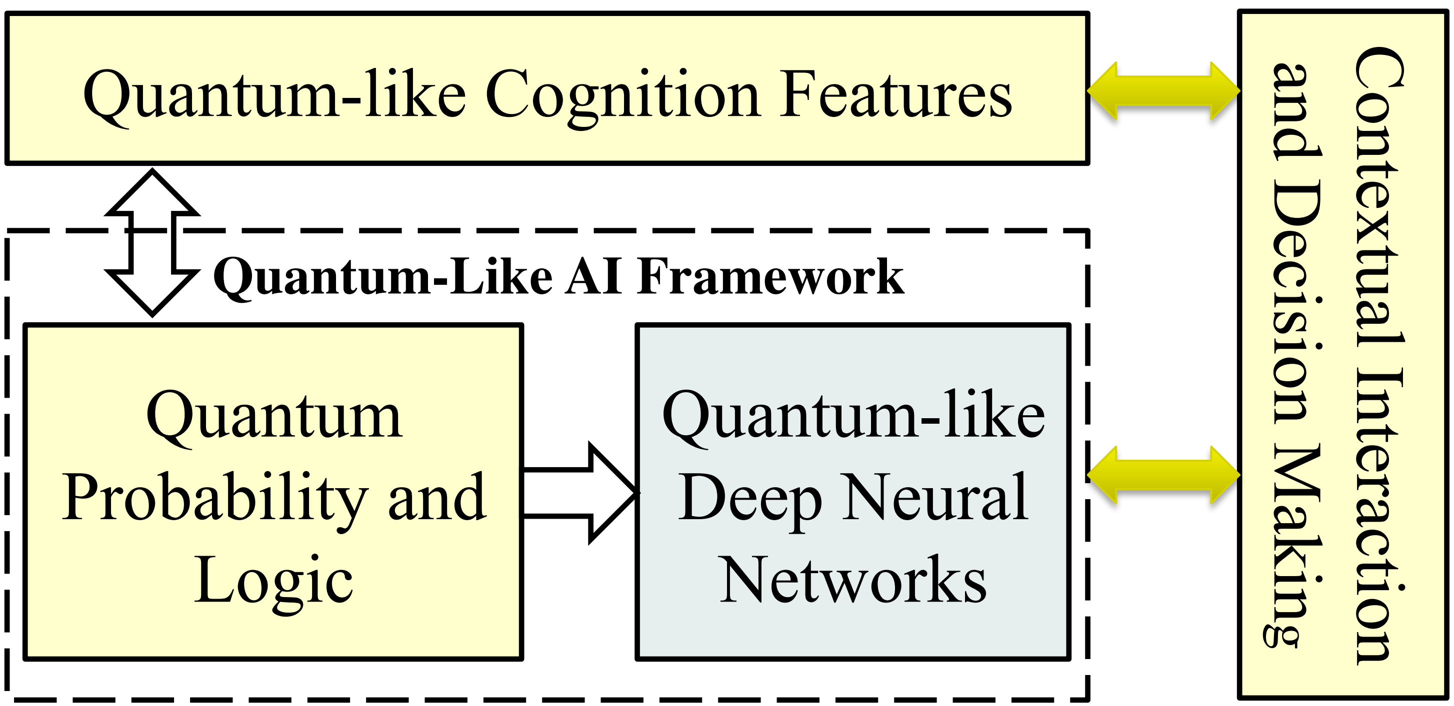
- Text Representation
  - Quantum language model (QLM) for ad-hoc retrieval
  - QLM for sentiment analysis
  - Adaptive QLM for session search
  - End-to-End QLM for question answering
  - Many-body Wave Function Inspired LM
  - Tensor Space LM
  - QLM for multimodal sentiment analysis
  - Tensor-based multimodal content fusion

More Quantum-inspired IR models

- Modelling user interaction and decision making
  - Quantum-interference inspired decision fusion model for multimodal sentiment analysis
  - Quantum-inspired conversational sentiment analysis model

A Broader Quantum-Cognitive Perspective for Artificial Intelligence

Future AI that is compatible with human cognitive processing



References

Modelling "quantumness" of user interaction and decision making in IR:

- Uprety, S., Dehdashti, S., Fell, L., Bruza, P.D., and **Song, D.** (2019). Modelling Dynamic Interactions Between Relevance Dimensions. The 5th ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR 2019). Best Full Paper Honorable Mention Award.
- Uprety, S., and **Song, D.** (2018). Investigating Order Effects in Multidimensional Relevance Judgment using Query Logs. The 4th ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR 2018), pp. 191-194.
- Uprety, S., and **Song, D.** (2018). Investigating Bell Inequalities for Multidimensional Relevance Judgments in Information Retrieval. Quantum Interaction (QI'2018).
- Gkoumas, D., Uprety, S., and **Song, D.** (2018). Investigating non-Classical Correlations Between Decision Fused Documents. Quantum Interaction (QI'2018).
- Uprety, S., Su, Y., **Song, D.,** Li, J. (2018). Modeling multidimensional user relevance in IR using vector spaces. ACM SIGIR 2018, short paper, pp. 993-996.
- Zhang, Y., **Song, D.,** Zhang, P., Wang, P., Li, J., Wang, B. (2018). A Quantum-Inspired Multimodal Sentiment Analysis Framework. Theoretical Computer Science. 752, pp. 21-40.
- Wang, P., Wang, T., Hou, Y., and **Song, D.** (2018). Modeling Relevance Judgement Inspired by Quantum Weak Measurement. The 40th European Conference on Information Retrieval (ECIR 2018), pp. 424-436.

References

Complex Semantic Hilbert Space for Text Representation:

- Qiuchi Li, Benyou Wang and Massimo Melucci. (2019). "CNM: An Interpretable Complex -valued Network for Matching." 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2019). Best Explanable Paper Award.
- Wang, B., Li, Q., Melucci, M., Song, D. (2019). Semantic Hilbert Space for Text Representation Learning. WWW2019, pp.3293-3299. May 13-17, San Francisco, California, USA.
- Li, Q., Uprety, S., Wang, B., Song, D. (2018). Quantum-inspired Complex Word Embedding. ACL Workshop on Representation Learning for NLP (Rep4NLP@ACL 2018), pp. 50-57. Melbourne, Australia, July 20, 2018.

References

Quantum Language Models:

- Zhang, L., Zhang, P., Ma, X., Gu, S., Su, Z., **Song, D.** (2019). A Generalized Language Model in Tensor Space. The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI 2019), pp. 7450-7458.
- Zhang, P., Su, Z., Zhang, L., Wang, B., **Song, D.** (2018). A Quantum Many-body Wave Function Inspired Language Modeling Approach. ACM CIKM2018, pp. 1303-1312. Zhang, P., Niu, J., Su, Z., Wang, B., **Song, D.** (2018). End-to-End Quantum-like Language Models with Application to Question Answering. AAAI 2018, pp. 5666-5673.
- Li, J., Zhang, P., **Song, D.,** and Hou, Y. (2016). An Adaptive Contextual Quantum Language Model. Physica A: Statistical Mechanics and its Applications. Vol. 456, pp. 51-67.
- Li, Q., Li, J., Zhang, P., **Song, D.** (2015). Modeling Multi-query Retrieval Tasks Using Density Matrix Transformation. ACM SIGIR2015, pp. 871-874.
- Xie, M., Hou, Y., Zhang, P., Li., J, Li, W., Song, D. (2015). Modeling Quantum Entanglements in Quantum Language Models. IJCAI2015, pp. 1362-1368.

References

Quantum Language Models for Sentiment Analysis:

- Zhang, Y., **Song, D.,** Zhang, P., Li, Xiang, Wang, P. (2019). A quantum-inspired sentiment representation model for twitter sentiment analysis. Applied Intelligence.49(8): 3093-3108.
- Zhang, Y., Li, Q., **Song, D.,** Zhang, P., Wang, P. (2019). Quantum-inspired Interactive Networks for Conversational Sentiment Analysis. The 28th International Joint Conference on Artificial Intelligence (IJCAI2019), pp. 5436-5442. 10-16 August 2019, Macau, China.
- Zhang, Y., **Song, D.,** Zhang, P., Wang, P., Li, J., Wang, B. (2018). A Quantum-Inspired Multimodal Sentiment Analysis Framework. Theoretical Computer Science. 752, pp. 21-40.
- Zhang, Y., **Song, D.,** Li, X. and Zhang, P. (2018). Unsupervised Sentiment Analysis of Twitter Posts Using Density Matrix Representation. The 40th European Conference on Information Retrieval (ECIR 2018), pp. 316-329. 26-29 March 2018, Grenoble, France.

Thank you!