UNIVERSITY OF CAMBRIDGE

Representations of Meaning

Yarin Gal

yg279@cam.ac.uk

22nd February 2015

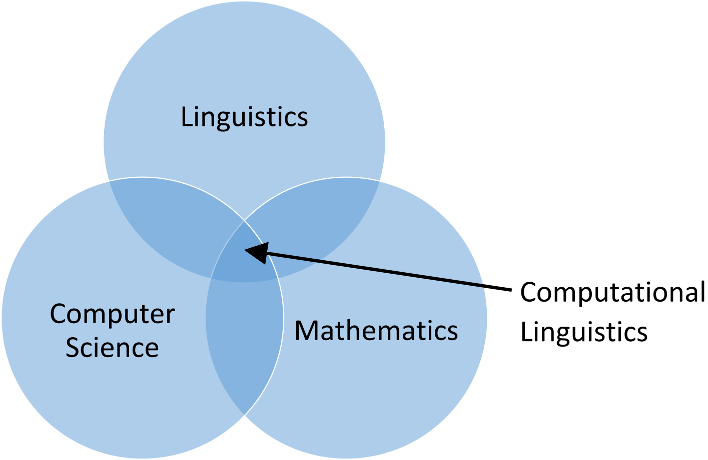
Logic in Ancient Greece ($\sim$ 300 BC)

Aristotle's famous example:

- All men are mortal
- Socrates is a man
- Therefore, Socrates is mortal

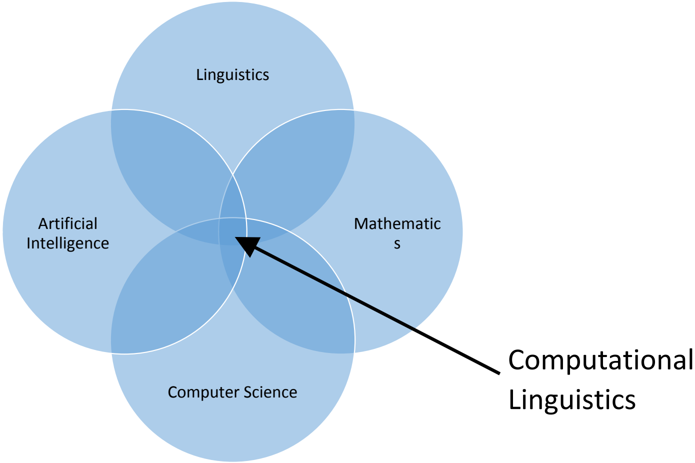
Computations Linguistics (1950s)

UNIVERSITY OF CAMBRIDGE

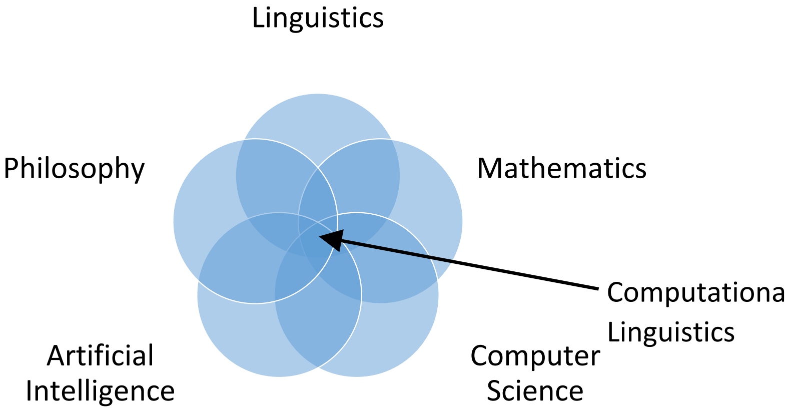


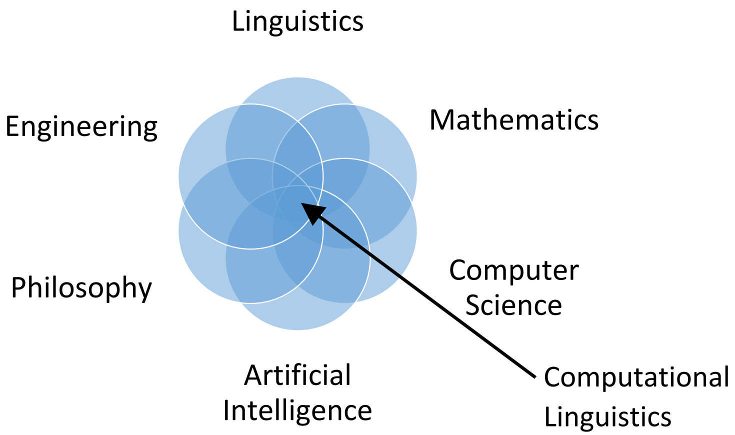
Computations Linguistics (1950s)

UNIVERSITY OF CAMBRIDGE

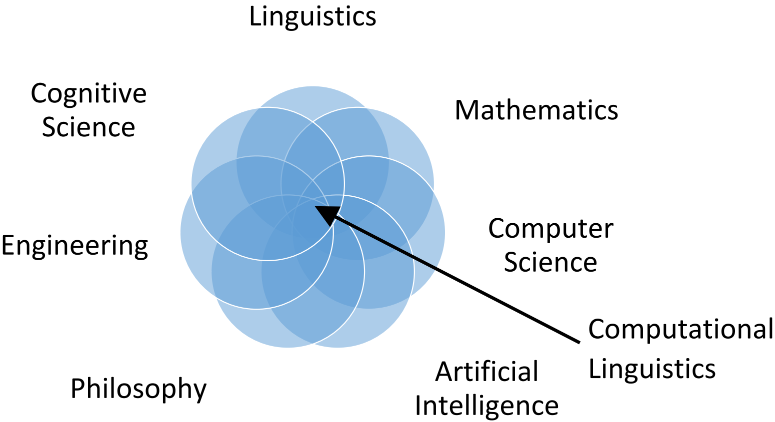


UNIVERSITY OF CAMBRIDGE




Linguistics
Engineering
Mathematics
Philosophy
Computer Science
Artificial Intelligence
Computational Linguistics

UNIVERSITY OF CAMBRIDGE



Gentzen Sequent Calculus (1934)

UNIVERSITY OF CAMBRIDGE

- $A_1, ..., A_n \vdash A$: A can be proved from assumptions $A_1, ..., A_n$
- Γ, Δ: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
- Identity and cut rule—

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

- Conjunction—

$$\frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} \wedge R \qquad \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C} \wedge L$$

- Implication—

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \to R \qquad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \to B, \Delta \vdash C} \to L$$

- Structural rules—

$$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}\text{Exchange} \qquad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\text{Contraction} \qquad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}\text{Weakening}$$

- **Equivalent to the Natural Deduction system.**

(based on notes by Samson Abramsky)

Gentzen Sequent Calculus (1934)

UNIVERSITY OF CAMBRIDGE

- $A_1, ..., A_n \vdash A$: A can be proved from assumptions $A_1, ..., A_n$
- Γ, Δ: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
- Identity and cut rule—

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

- Conjunction—

$$\frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} \wedge R \qquad \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C} \wedge L$$

- Implication—

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \to R \qquad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \to B, \Delta \vdash C} \to L$$

- Structural rules—

$$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}\text{Exchange} \qquad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\text{Contraction} \qquad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}\text{Weakening}$$

- **Equivalent to the Natural Deduction system.**

(based on notes by Samson Abramsky)

- $A_1, ..., A_n \vdash A$: A can be proved from assumptions $A_1, ..., A_n$
- Γ, Δ: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
- Identity and cut rule—

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

- Conjunction—

$$\frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} \wedge R \qquad \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C} \wedge L$$

- Implication—

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \rightarrow R \qquad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \rightarrow B, \Delta \vdash C} \rightarrow L$$

- Structural rules—

$$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}\text{Exchange} \qquad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\text{Contraction} \qquad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}\text{Weakening}$$

- **Equivalent to the Natural Deduction system.**

(based on notes by Samson Abramsky)

- $A_1, ..., A_n \vdash A$: A can be proved from assumptions $A_1, ..., A_n$
- Γ, Δ: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
- Identity and cut rule—

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

- Conjunction—

$$\frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} \wedge R \qquad \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C} \wedge L$$

- Implication—

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \rightarrow R \qquad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \rightarrow B, \Delta \vdash C} \rightarrow L$$

- Structural rules—

$$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}\text{Exchange} \qquad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\text{Contraction} \qquad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}\text{Weakening}$$

- **Equivalent to the Natural Deduction system.**

(based on notes by Samson Abramsky)

- $A_1, ..., A_n \vdash A$: A can be proved from assumptions $A_1, ..., A_n$
- Γ, Δ: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
- Identity and cut rule—

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

- Conjunction—

$$\frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} \wedge R \qquad \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C} \wedge L$$

- Implication—

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \to R \qquad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \to B, \Delta \vdash C} \to L$$

- Structural rules—

$$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}\text{Exchange} \qquad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\text{Contraction} \qquad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}\text{Weakening}$$

- **Equivalent to the Natural Deduction system.**

(based on notes by Samson Abramsky)

Gentzen Sequent Calculus (1934)

UNIVERSITY OF CAMBRIDGE

- $A_1, ..., A_n \vdash A$: A can be proved from assumptions $A_1, ..., A_n$
- Γ, Δ: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
- Identity and cut rule—

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

- Conjunction—

$$\frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} \wedge R \qquad \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C} \wedge L$$

- Implication—

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \to R \qquad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \to B, \Delta \vdash C} \to L$$

- Structural rules—

$$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}\text{Exchange} \qquad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\text{Contraction} \qquad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}\text{Weakening}$$

- **Equivalent to the Natural Deduction system.**

(based on notes by Samson Abramsky)

Gentzen Sequent Calculus (1934)

UNIVERSITY OF CAMBRIDGE

- $A_1, ..., A_n \vdash A$: A can be proved from assumptions $A_1, ..., A_n$
- Γ, Δ: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
- Identity and cut rule—

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

- Conjunction—

$$\frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} \wedge R \qquad \frac{\Gamma, A, B \vdash C}{\Gamma, A \wedge B \vdash C} \wedge L$$

- Implication—

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \to R \qquad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \to B, \Delta \vdash C} \to L$$

- Structural rules—

$$\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C}\text{Exchange} \qquad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\text{Contraction} \qquad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}\text{Weakening}$$

- **Equivalent to the Natural Deduction system.**

(based on notes by Samson Abramsky)

Gentzen Sequent Calculus – Example

- The sentences “All men are mortal” and “Socrates is a man” are mapped to

$$\text{man} \vdash \text{mortal}; \quad \text{Socrates} \vdash \text{man}.$$

  Using the cut rule we get

$$\frac{\text{man} \vdash \text{mortal}; \quad \text{Socrates} \vdash \text{man}}{\text{Socrates} \vdash \text{mortal}}\text{Cut}$$

  And infer that Socrates is mortal.

- But, what about

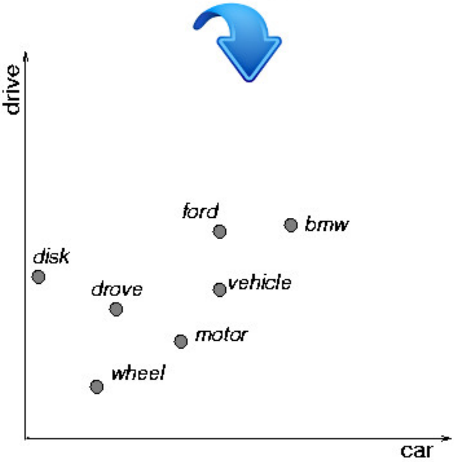
$$\text{went} \rightarrow \text{pub} \overset{?}{\vdash} \text{went} \rightarrow \text{bar}$$

You shall know a word by the company it keeps.

–Firth, J. R. 1957:11

Vector space over $\mathbb{R}$

$$
\begin{array}{l}
\\
\textit{term } 1 \\
\textit{term } 2 \\
\\
\textit{term } 300{,}000
\end{array}
\begin{array}{c}
\begin{array}{cccccccc}
d_1 & d_2 & d_3 & d_4 & d_5 & d_6 & & d_{3,000,000,000}
\end{array} \\
\begin{pmatrix}
0 & 0 & 1 & 0 & 1 & 0 & \cdots & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0
\end{pmatrix}
\end{array}
$$

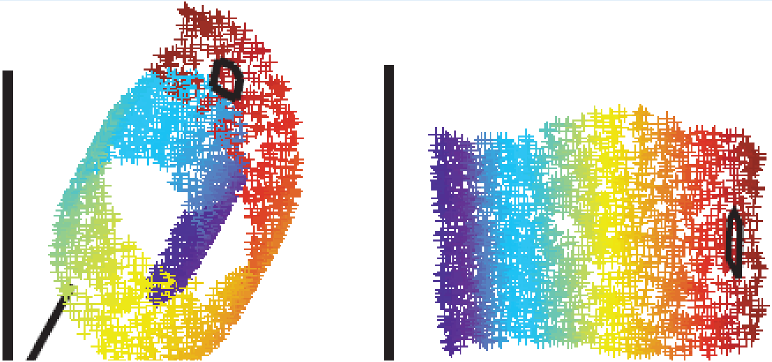


The Johnson-Lindenstrauss Lemma

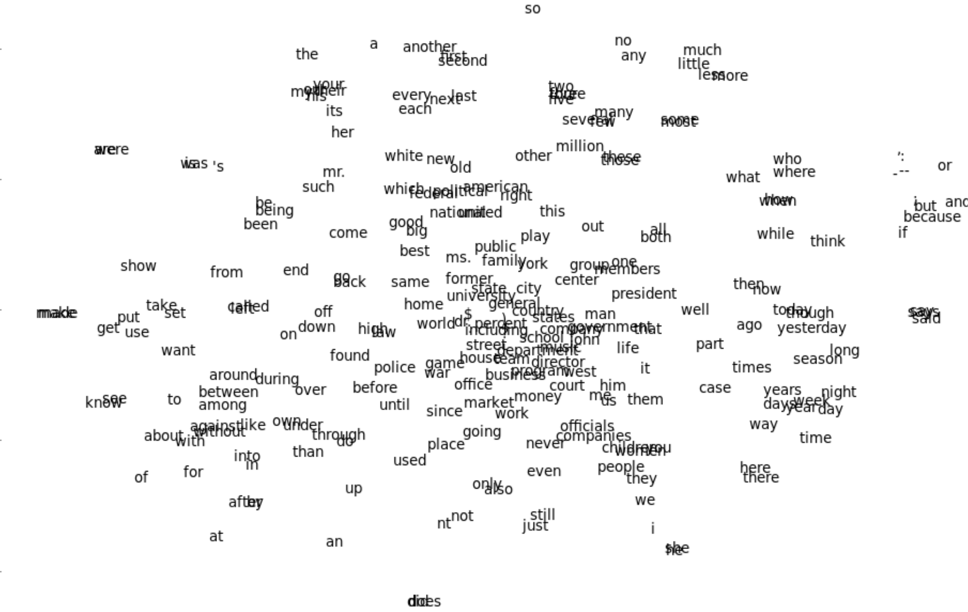
For any $0 < \epsilon < 1/2$ and any integer $m > 4$, let $k = \frac{20 \log m}{\epsilon^2}$. Then, for any set V of m points in $\mathbb{R}^N$, $\exists f : \mathbb{R}^N \to \mathbb{R}^k$ s.t. $\forall u, v \in V$:

$$(1\epsilon)||uv||^2 \leq ||f(u)f(v)||^2 \leq (1+\epsilon)||uv||^2$$

e.g. $f(x) = \frac{1}{\sqrt{k}} Ax$ with $A_{ij} \in \mathcal{N}(0,1)$.



Vector space over $\mathbb{R}$ – Example



A category **C** consists of:

- A set *ob*(**C**) of objects,
- A set *hom*(**C**) of morphisms, or arrows. Each arrow f has a source object A and target object B,
- An identity arrow id_A for every object A.
- An associative composition operation between arrows $\circ$.

Composing $f : A \rightarrow B$ and $g : B \rightarrow C$ gives $g \circ f$ from A to C.

A **symmetric monoidal closed category** is

- a category **C**
- equipped with a symmetric associative bifunctor $\otimes : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ called the tensor product
- and an object I called the identity object[1].
- For all objects A and B there is an object

  $$A \multimap B$$

  and an arrow

  $$ev_{A,B} : (A \multimap B) \otimes A \to B.$$

  For every arrow $f : C \otimes A \to B$, there is a unique arrow $\Lambda(f) : C \to (A \multimap B)$ such that $ev_{A,B} \circ (\Lambda(f) \otimes \mathrm{id}_A) = f$.

[1] The symmetry, associativity and identity are define through natural isomorphisms.

Symmetric monoidal closed categories

The category of **finite dimensional vector spaces** is a symmetric monoidal closed category.

- The tensor product is the tensor product of vector spaces,
- and $A \multimap B$ is the vector space of linear maps.

Gentzen sequent calculus without the Contraction and Weakening rules also corresponds to symmetric monoidal closed categories.

- Known as **linear logic**,
- A “resource-sensitive” logic[2],
- The tensor product is the conjunction $\wedge$,
- and $A \multimap B$ is the implication $\rightarrow$.

For example, the identity and cut rule —

$$\frac{}{A \vdash A}\text{Id} \qquad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}\text{Cut}$$

are equivalent to

$$\frac{}{\text{Id}_A : A \rightarrow A} \qquad \frac{f : \Gamma \rightarrow A; \quad g : A \otimes \Delta \rightarrow B}{g \circ (f \otimes \text{Id}_\Delta) : \Gamma \otimes \Delta \rightarrow B}$$

[2] It is possible to recover the expressive power of standard Gentzen sequent calculus with the addition of some connectives.

The Compositional Distributional representation:

- Let $N = \mathbb{R}$ be a *noun vector space*, with nouns represented as vectors,
- Let $V = \mathbb{R}^3$ be a *verb vector space*, with verbs represented as third-order tensors,
- Let S be a *sentence vector space*, representing sentences as the tensor product $N \otimes V \otimes N$,
- Finally let T be a *truth value vector space*, a lower dimensional vector space to which we project products.

For example,

- "Dogs" and "Cats" are represented as vectors d and c
- "chase" is represented as $T \in \mathbb{R}^3$
- "Dogs chase cats" is represented as $d \otimes T \otimes c$

For simplicity, we use a **binary noun vector space**, and T is a binary matrix:

- The first dimension of N is "*likes chasing small fluffy animals*" and the 2nd and 3rd dimensions are "*is small*" and "*is fluffy*",
- A cat is represented as small and fluffy $c = [0, 1, 1]$, a dog likes to chase small and fluffy animals $d = [1, 0, 0]$,
- Tensor "chase" preserves vectors from the left that have the "likes chasing small fluffy animals" property and vectors from the right that have the "is small and is fluffy properties,

$$T = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Compositional Distributional – Example

UNIVERSITY OF CAMBRIDGE

$$c = [0, 1, 1], \quad d = [1, 0, 0]$$

$$T = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad t = dTc^T$$

- The proposition “dogs chase cats” is mapped to $dTc^T = 1$ – a high “truthness” value.
- The proposition “cats chase dogs” is mapped to $cTd^T = 0$ – low truthness.

UNIVERSITY OF CAMBRIDGE

The talk was based on [Gal, 2013].

Gal, Y. (2013).
Semantics, modelling, and the problem of representation of meaning – a brief survey of recent literature.
Technical report, University of Cambridge.