Representations of Meaning

Yarin Gal

yg279@cam.ac.uk

22nd February 2015
Aristotle’s famous example:

- All men are mortal
- Socrates is a man
- Therefore, Socrates is mortal
Computations Linguistics (1950s)

Linguistics

Computer Science

Mathematics

Computational Linguistics
Computations Linguistics (1950s)

Linguistics
Mathematics
Computer Science
Artificial Intelligence
Computational Linguistics
Computations Linguistics (1950s)
A_1, ..., A_n \vdash A: A can be proved from assumptions A_1, ..., A_n

\Gamma, \Delta: finite lists of formulas, define \Gamma, A := \Gamma \cup \{A\}

Identity and cut rule—
\[
\begin{align*}
\frac{}{A \vdash A} \text{Id} & \\
\frac{\Gamma \vdash A; \ A, \Delta \vdash B}{\Gamma, \Delta \vdash B} \text{Cut}
\end{align*}
\]

Conjunction—
\[
\begin{align*}
\frac{\Gamma \vdash A; \ \Delta \vdash B}{\Gamma, \Delta \vdash A \land B} & \land R \\
\frac{\Gamma, A, B \vdash C}{\Gamma, A \land B \vdash C} \land L
\end{align*}
\]

Implication—
\[
\begin{align*}
\frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \rightarrow R & \\
\frac{\Gamma \vdash A; \ B, \Delta \vdash C}{\Gamma, A \rightarrow B, \Delta \vdash C} \rightarrow L
\end{align*}
\]

Structural rules—
\[
\begin{align*}
\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C} \text{Exchange} & \\
\frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B} \text{Contraction} & \\
\frac{\Gamma \vdash B}{\Gamma, A \vdash B} \text{Weakening}
\end{align*}
\]

Equivalent to the Natural Deduction system.

(based on notes by Samson Abramsky)
Gentzen Sequent Calculus (1934)

▶ $A_1, ..., A_n \vdash A$: $A$ can be proved from assumptions $A_1, ..., A_n$
▶ $\Gamma, \Delta$: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$

▶ Identity and cut rule—

\[
\begin{align*}
& \quad \Gamma \vdash A \quad \text{Id} \\
& \quad \Gamma, A, \Delta \vdash B \quad \text{Cut}
\end{align*}
\]

▶ Conjunction—

\[
\begin{align*}
& \quad \Gamma, \Delta \vdash A \\
& \quad \Delta \vdash B \\
\hline
& \quad \Gamma, \Delta \vdash A \land B \\
\quad \wedge R
\end{align*}
\]

\[
\begin{align*}
& \quad \Gamma, A, B \vdash C \\
\hline
& \quad \Gamma, A \land B \vdash C \\
\quad \wedge L
\end{align*}
\]

▶ Implication—

\[
\begin{align*}
& \quad \Gamma, A \vdash B \\
\hline
& \quad \Gamma \vdash A \rightarrow B \\
\quad \rightarrow R
\end{align*}
\]

\[
\begin{align*}
& \quad \Gamma \vdash A \\
& \quad B, \Delta \vdash C \\
\hline
& \quad \Gamma, A \rightarrow B, \Delta \vdash C \\
\quad \rightarrow L
\end{align*}
\]

▶ Structural rules—

\[
\begin{align*}
& \quad \Gamma, A, B, \Delta \vdash C \\
\hline
& \quad \Gamma, B, A, \Delta \vdash C \\
\quad \text{Exchange}
\end{align*}
\]

\[
\begin{align*}
& \quad \Gamma, A, A \vdash B \\
\hline
& \quad \Gamma, A \vdash B \\
\quad \text{Contraction}
\end{align*}
\]

\[
\begin{align*}
& \quad \Gamma \vdash B \\
\hline
& \quad \Gamma, A \vdash B \\
\quad \text{Weakening}
\end{align*}
\]

▶ Equivalent to the Natural Deduction system.

(based on notes by Samson Abramsky)
Gentzen Sequent Calculus (1934)

- $A_1, ..., A_n \vdash A$: $A$ can be proved from assumptions $A_1, ..., A_n$
- $\Gamma, \Delta$: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{ A \}$
- Identity and cut rule—
  \[
  \frac{}{A \vdash A} \text{Id} \quad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B} \text{Cut}
  \]
- Conjunction—
  \[
  \frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \land B} \land R \quad \frac{\Gamma, A, B \vdash C}{\Gamma, A \land B \vdash C} \land L
  \]
- Implication—
  \[
  \frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \rightarrow R \quad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \rightarrow B, \Delta \vdash C} \rightarrow L
  \]
- Structural rules—
  \[
  \frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C} \text{Exchange} \quad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B} \text{Contraction} \quad \frac{\Gamma \vdash B}{\Gamma, A \vdash B} \text{Weakening}
  \]
- Equivalent to the Natural Deduction system.

(based on notes by Samson Abramsky)
A_1, \ldots, A_n \vdash A: A can be proved from assumptions A_1, \ldots, A_n

Γ, Δ: finite lists of formulas, define Γ, A := Γ \cup \{A\}

Identity and cut rule—

\[
\begin{align*}
\frac{}{A \vdash A} & \text{Id} \\
\frac{Γ \vdash A; A, Δ \vdash B}{Γ, Δ \vdash B} & \text{Cut}
\end{align*}
\]

Conjunction—

\[
\begin{align*}
\frac{Γ \vdash A; Δ \vdash B}{Γ, Δ \vdash A \land B} & \land R \\
\frac{Γ, A, B \vdash C}{Γ, A \land B \vdash C} & \land L
\end{align*}
\]

Implication—

\[
\begin{align*}
\frac{Γ, A \vdash B}{Γ \vdash A \rightarrow B} & \rightarrow R \\
\frac{Γ \vdash A; B, Δ \vdash C}{Γ, A \rightarrow B, Δ \vdash C} & \rightarrow L
\end{align*}
\]

Structural rules—

\[
\begin{align*}
\frac{Γ, A, B, Δ \vdash C}{Γ, B, A, Δ \vdash C} & \text{Exchange} \\
\frac{Γ, A, A \vdash B}{Γ, A \vdash B} & \text{Contraction} \\
\frac{Γ \vdash B}{Γ, A \vdash B} & \text{Weakening}
\end{align*}
\]

Equivalent to the Natural Deduction system.

(based on notes by Samson Abramsky)
▶ $A_1, \ldots, A_n \vdash A$: $A$ can be proved from assumptions $A_1, \ldots, A_n$
▶ $\Gamma, \Delta$: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{A\}$
▶ Identity and cut rule—
\[
\frac{}{A \vdash A} \quad \text{Id} \quad \frac{\Gamma \vdash A; \ A, \Delta \vdash B}{\Gamma, \Delta \vdash B} \quad \text{Cut}
\]
▶ Conjunction—
\[
\frac{\Gamma \vdash A; \ \Delta \vdash B}{\Gamma, \Delta \vdash A \land B} \quad \land R \quad \frac{\Gamma, A, B \vdash C}{\Gamma, A \land B \vdash C} \quad \land L
\]
▶ Implication—
\[
\frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \quad \rightarrow R \quad \frac{\Gamma \vdash A; \ B, \Delta \vdash C}{\Gamma, A \rightarrow B, \Delta \vdash C} \quad \rightarrow L
\]
▶ Structural rules—
\[
\frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C} \quad \text{Exchange} \quad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B} \quad \text{Contraction} \quad \frac{\Gamma \vdash B}{\Gamma, A \vdash B} \quad \text{Weakening}
\]
▶ Equivalent to the Natural Deduction system.

(based on notes by Samson Abramsky)
Gentzen Sequent Calculus (1934)

- \( A_1, \ldots, A_n \vdash A \): \( A \) can be proved from assumptions \( A_1, \ldots, A_n \)
- \( \Gamma, \Delta \): finite lists of formulas, define \( \Gamma, A := \Gamma \cup \{ A \} \)
- Identity and cut rule—
  \[
  \frac{}{A \vdash A} \text{Id} \quad \frac{\Gamma \vdash A; A, \Delta \vdash B}{\Gamma, \Delta \vdash B} \text{Cut}
  \]
- Conjunction—
  \[
  \frac{\Gamma \vdash A; \Delta \vdash B}{\Gamma, \Delta \vdash A \land B} \land R \quad \frac{\Gamma, A, B \vdash C}{\Gamma, A \land B \vdash C} \land L
  \]
- Implication—
  \[
  \frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \rightarrow R \quad \frac{\Gamma \vdash A; B, \Delta \vdash C}{\Gamma, A \rightarrow B, \Delta \vdash C} \rightarrow L
  \]
- Structural rules—
  \[
  \frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C} \text{Exchange} \quad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B} \text{Contraction} \quad \frac{\Gamma \vdash B}{\Gamma, A \vdash B} \text{Weakening}
  \]

Equivalent to the Natural Deduction system.
(based on notes by Samson Abramsky)
Gentzen Sequent Calculus (1934)

- $A_1, ..., A_n \vdash A$: $A$ can be proved from assumptions $A_1, ..., A_n$
- $\Gamma, \Delta$: finite lists of formulas, define $\Gamma, A := \Gamma \cup \{ A \}$
- Identity and cut rule—
  \[
  \frac{}{A \vdash A} \quad \frac{\Gamma \vdash A; \quad A, \Delta \vdash B}{\Gamma, \Delta \vdash B}
  \]

- Conjunction—
  \[
  \frac{\Gamma \vdash A; \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \land B} \quad \frac{\Gamma, A, B \vdash C}{\Gamma, A \land B \vdash C}
  \]

- Implication—
  \[
  \frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \quad \frac{\Gamma \vdash A; \quad B, \Delta \vdash C}{\Gamma, A \rightarrow B, \Delta \vdash C}
  \]

- Structural rules—
  \[
  \frac{\Gamma, A, B, \Delta \vdash C}{\Gamma, B, A, \Delta \vdash C} \quad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B} \quad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}
  \]

- Equivalent to the Natural Deduction system.

(based on notes by Samson Abramsky)
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

\[ \frac{\text{man} \vdash \text{mortal}; \quad \text{Socrates} \vdash \text{man}}{\text{Socrates} \vdash \text{mortal}} \]

And infer that Socrates is mortal.

But, what about

\[ \text{went} \rightarrow \text{pub} \vdash \text{went} \rightarrow \text{bar} \]
You shall know a word by the company it keeps.

–Firth, J. R. 1957:11
Vector space over \( \mathbb{R} \)
Dimensionality reduction

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 $\mathbb{R}^N$, $\exists f : \mathbb{R}^N \rightarrow \mathbb{R}^k$ 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 $\mathbf{C}$ consists of:

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

Composing $f : A \to B$ and $g : B \to C$ gives $g \circ f$ from $A$ to $C$. 
A symmetric monoidal closed category is

- a category $\mathbf{C}$
- 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 $A \rightarrow B$

and an arrow

$$\text{ev}_{A,B} : (A \rightarrow B) \otimes A \to B.$$ 

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

$^1$The symmetry, associativity and identity are define through natural isomorphisms.
The category of **finite dimensional vector spaces** is a symmetric monoidal closed category.

- The tensor product is the tensor product of vector spaces,
- and $A \leadsto B$ is the vector space of linear maps.
Symmetric monoidal closed categories

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 \(\land\),
- and \(A \rightarrow B\) is the implication \(\rightarrow\).

For example, the identity and cut rule —

\[
\frac{\Gamma \vdash A; \ A, \Delta \vdash B}{\Gamma, \Delta \vdash B} \quad \text{Cut}
\]

\[
\frac{\Gamma \vdash A; \ A, \Delta \vdash B}{\Gamma, \Delta \vdash B} \quad \text{Cut}
\]

are equivalent to

\[
\frac{\text{Id}_A : A \rightarrow A}{f : \Gamma \rightarrow A; \ g : A \otimes \Delta \rightarrow B}
\]

\[
\frac{\text{Id}_A : A \rightarrow A}{f : \Gamma \rightarrow A; \ g : A \otimes \Delta \rightarrow B}
\]

\[
\frac{g \circ (f \otimes \text{Id}_\Delta) : \Gamma \otimes \Delta \rightarrow B}{g \circ (f \otimes \text{Id}_\Delta) : \Gamma \otimes \Delta \rightarrow B}
\]

\(^2\)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 \textit{noun vector space}, with nouns represented as vectors,
- Let $V = \mathbb{R}^3$ be a \textit{verb vector space}, with verbs represented as third-order tensors,
- Let $S$ be a \textit{sentence vector space}, representing sentences as the tensor product $N \otimes V \otimes N$,
- Finally let $T$ be a \textit{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,

\[
T = \begin{bmatrix}
0 & 1 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}.
\]
\[ c = [0, 1, 1], \quad d = [1, 0, 0] \]

\[
T = \begin{bmatrix}
0 & 1 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}, \quad t = dTC^T
\]

- 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.
Thank you for listening

The talk was based on [Gal, 2013].