# LING5702: Lecture Notes 2 Models of Thought

#### **Contents**

2.1	Decision theory [von Neumann & Morgenstern, 1944]	1
2.2	Typed lambda calculus [Church, 1940]	2
2.3	Generalized quantifiers [Barwise & Cooper, 1981]	3
2.4	Complex (nested) propositions	5
2.5	Example: grid-world navigation	5
2.6	Extra: quantifiers over substances	5
2.7	Extra: propositions about arbitrary probabilities	7
2.8	Extra: intensions (propositions about propositions)	7
2.9	No need for other operators	9
2.10	Review	9

Language seems similar to thought, but we can distinguish them. Let's start with thought.

# 2.1 Decision theory [von Neumann & Morgenstern, 1944]

We model thought as a **decision process**: we choose actions to maximize **average expected utility**. (There's lots of other thought: enjoying, reminiscing, etc., but this is about what helps us survive.)

A decision process assumes a set of **plans** p, a **world model** m and a **reward** R(m). For example:

- p may be a plan to walk to a hill,
- m may include our location and the knowledge that a step may take us closer to a goal,
- R(m) may be the prestige we get from reaching the hill.

Average (over time t) expected reward for a plan p is a sum over outcomes o of actions a in p:

$$\mathsf{AEU}(t, m, p) = \mathsf{R}(m) \cdot \begin{cases} \mathsf{if} \ \exists_a \ p \land m \to a : \\ \mathsf{otherwise:} \end{cases} \underbrace{\sum_o \mathsf{P}(o \mid m \land a)}_{\mathsf{new} \ m \ \mathsf{at} \ \mathsf{next} \ t} \underbrace{\mathsf{Tepeat} \ \mathsf{with} \ m, a \ \mathsf{and} \ o \ \mathsf{as} \ \mathsf{new} \ m}_{\mathsf{new} \ m \ \mathsf{at} \ \mathsf{next} \ t}$$

(It assumes the plan p is perfectly specific: at most one next action a for each possible model m.)

For example, if the goal hill is one step away, we get a reward in one step, so AEU(t, m, p) = 1.

But if it's muddy and we slip half the time and don't go anywhere, then:

$$\mathsf{AEU}(t, m, p) = \begin{cases} .5 \text{ (no slip)} & \times \frac{1}{1} \text{ (arrive in 1 step)} \\ +.5 \text{ (slip)} & \times \begin{cases} .5 \text{ (no slip)} & \times \frac{1}{2} \text{ (arrive in 2 steps)} \\ +.5 \text{ (slip)} & \times \begin{cases} .5 \text{ (no slip)} & \times \frac{1}{3} \text{ (arrive in 3 steps)} \\ +.5 \text{ (slip)} & \times \end{cases} \\ \approx .7 \end{cases}$$

So if we have two plans (clear path and muddy path) and we know mud slows us, we can avoid it.

We speculate language provides us an advantage by letting us **share** plans and world knowledge. So what are these plans and world knowledge?

# 2.2 Typed lambda calculus [Church, 1940]

We use **typed lambda calculus** to model complex ideas that make up plans and world knowledge. We'll write  $[\![\ldots]\!]_m$  for the **denotation** or **interpretation** of expression '...' in world model m. (World models may be underspecified, so they are **sets of possible worlds**:  $[\![\ldots]\!]_m = \forall_{w \in m} [\![\ldots]\!]_w$ .)

Typed lambda calculus expressions have the following types:

- 1. **entity terms**: references to things that can be predicated over, like people and places; They each denote an **entity** in m if one exists, e.g.  $[MyHill]_m = Hill1$ , or are ill-formed if not.
- 2. **propositions**: things that can be true or false, like the proposition that it is sunny; They each denote a **truth value** as evaluated in m, e.g. [[ItsSunny]] $_m =$ True.
- 3. **functions** from any type as input to any type as output (including other functions). They denote **sets** of input-output pairs, e.g.  $[Muddy]_m = {\langle Hill1, True \rangle, \langle Hill2, False \rangle, ...}$ .

Typed lambda calculus expressions are constructed using the following **rules**:

```
1. applications of functions f to arguments x to get outputs y = f(x), like Muddy MyHill function entity term
```

This retrieves the (unique) output:  $[\![f\,x]\!]_m = h$  such that  $\langle [\![x]\!]_m, h \rangle \in [\![f]\!]_m$ , e.g. True.

function from entity term x to proposition

2. **abstractions** over argument variables x to get functions  $f = \lambda_x ... x...$ , like  $\lambda_x$  Muddy x proposition (truth value)

This creates a set of pairs  $[\![\lambda_x ...x...]\!]_m = \{\langle x, [\![...x...]\!]_m \rangle \mid x \in m\}$  (constrained to x's type). If truth output, use set of input:  $[\![\lambda_x ...x...]\!]_m = \{x \mid x \in m, [\![...x...]\!]_m\}$ , so  $[\![Muddy]\!]_m = \{Hill1\}$ .

The most common functions we need are:

predicate entity pred.ent. ent.

1. **predicates**: e.g. Person x or At x y, which map entity terms to truth values (propositions)

For extra entities, use functions as output:  $[At]_m = {\langle Me, {\langle Hill1, True \rangle, \langle Hill2, False \rangle, ...} \rangle, ...}$ 

output is another function with second entity as input!

2. **conjunctions**: e.g. Person  $x \wedge \text{At } x \text{ MyHill}$ , which map truth values to truth values:

 $\llbracket \varphi \wedge \psi \rrbracket_m \iff \llbracket \wedge \varphi \psi \rrbracket_m$ (typically used with 'infix' notation: function in middle)

3. **generalized quantifiers**: e.g. All  $(\lambda_x | \text{Person } x)$   $(\lambda_x | \text{At } x | \text{MyHill})$ , map sets to truth values

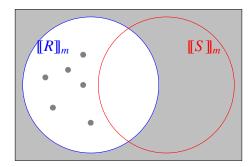
#### Practice: notation

Using the predicates  $\operatorname{Dog} x$ , which means x is a dog, and Mammal x, which means x is a mammal, write a typed lambda calculus expression stating that all dogs are mammals.

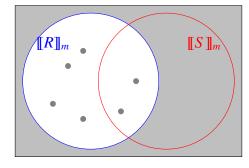
#### Generalized quantifiers [Barwise & Cooper, 1981] 2.3

Generalized quantifiers compare denotations of intersections of restriction R and nuclear scope S:

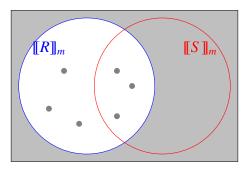
• [None RS]<sub>m</sub>  $\Leftrightarrow$  |[[R]]<sub>m</sub>  $\cap$  [[S]]<sub>m</sub>| = 0 — true if none of the R's are S's:



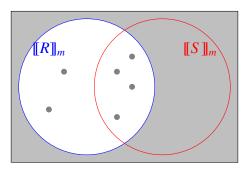
• [Some RS]]<sub>m</sub>  $\Leftrightarrow$  |[[R]]<sub>m</sub>  $\cap$  [[S]]<sub>m</sub>| > 0 — true if some of the R's are S's:



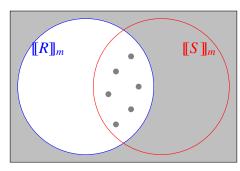
• [[Half RS]]<sub>m</sub>  $\Leftrightarrow$   $\frac{\|[R]_m \cap [S]\|_m}{\|[R]\|_m|} = 0.5$  — true if half of the R's are S's:



•  $[Most R S]_m \Leftrightarrow \frac{|[R]_m \cap [S]_m|}{|[R]_m|} > 0.5$  — true if most of the R's are S's:



• [[All R S]]<sub>m</sub>  $\Leftrightarrow \frac{\|[R]\|_m \cap [[S]]_m|}{\|[R]\|_m|} = 1.0$  — true if all of the R's are S's:



Note the similarity to diagrams of conditional probabilities from the previous lecture notes. Generalized quantifiers represent conditional probabilities  $P_m(S \mid R)$ , so we use them for reasoning! (Specifically, we use them to represent probabilistic outcome events o in our decision processes.)

# Practice: meaning

Given a world m of Shape entities (where Red and Square have their usual meanings):



what is the value of the following lambda calculus expression?

```
\llbracket Most (\lambda_x \text{ Shape } x \land Red x) (\lambda_x \text{ Square } x) \rrbracket_m
```

Practice: another meaning

Given the same set of shapes above, what is the value of the following lambda calculus expression?

```
\llbracket Most (\lambda_x \text{ Shape } x) (\lambda_x \text{ Square } x \land \text{Red } x) \rrbracket_m
```

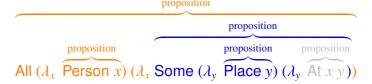
# 2.4 Complex (nested) propositions

Recall that quantifiers are propositions that can contain propositions:

Proposition

All 
$$(\lambda_x | \text{Person } x)$$
  $(\lambda_x | \text{At } x | \text{MyHill})$ 

This means we can stuff them inside each other like Russian dolls or turduckens:



This is called **nesting** or **scoping**.

# 2.5 Example: grid-world navigation

Now we can make a plan and world model that moves a person toward a hill in a grid world:

- 1. a **plan** p to move Me to MyHill (assuming the following predicates:
  - CurrentTime t is true for the most recent time point t in an AEU evaluation;
  - PrecedeOrEqual s t is true if time s precedes or is equal to t;
  - TryMoveToward t a x is true if a tries to move toward x at time t;

where 'All's iterate over entities, 'None's give conditions):

```
All (\lambda_t \text{ CurrentTime } t \land \text{None } (\lambda_s \text{ PrecedeOrEqual } 0 \text{ } s \land \text{ PrecedeOrEqual } s \text{ } t)
(\lambda_s \text{ At } s \text{ Me MyHill}))
(\lambda_t \text{ TryMoveToward } t \text{ Me MyHill})
```

(Here time '0' is when the plan is created – I have to reach the goal *after* that to succeed.)

When used in an AEU, this gives us our actions a – in this case: TryMoveToward predicates.

- 2. **world knowledge** *m* that moving through mud may fail (assuming the following predicates:
  - Adjacent x y is true if grid squares x and y are adjacent;
  - Aligned x y z is true if a grid square y lies on a line from x to z;
  - Muddy y and Clear y are true if grid square y is muddy or clear, respectively;
  - ConsecutiveTime t u is true if time t immediately precedes time u;

where 'Half's give probability cost):

```
All (\lambda_{t,a,z} \text{ TryMoveToward } t \ a \ z)
(\lambda_{t,a,z} \text{ All } (\lambda_x \text{ At } t \ a \ x)
(\lambda_x \text{ All } (\lambda_y \text{ Adjacent } x \ y \ \land \text{ Aligned } x \ y \ z \ \land \text{ Muddy } y)
(\lambda_y \text{ Half } (\lambda_u \text{ ConsecutiveTime } t \ u) \ (\lambda_u \text{ At } u \ a \ y) \ \land 
\text{Half } (\lambda_u \text{ ConsecutiveTime } t \ u) \ (\lambda_u \text{ At } u \ a \ x))))
```

and that moving through clear terrain always succeeds:

```
All (\lambda_{t,a,z} TryMoveToward t a z)
(\lambda_{t,a,z} All (\lambda_x At t a x)
(\lambda_x All (\lambda_y Adjacent x y \wedge Aligned x y z \wedge Clear y)
(\lambda_y All (\lambda_u ConsecutiveTime t u) (\lambda_u At u a y))))
```

When used in an AEU, this gives us our outcome events o – in this case: At predicates.

Dropping these plans and world models into the AEU function defines a rational decision process. Since they are *simple* and *work*, they are in some sense a 'natural' representation of complex ideas. We'll therefore use these expressions as the complex ideas that get communicated using language.

But note these look very different from how we might represent these ideas in natural language.

# 2.6 Extra: quantifiers over substances

Generalized quantifiers model substances as sets of infinitesimal (arbitrarily small) 'minimal parts':

```
Most (\lambda_v Contain MilkyWayGalaxy v) (\lambda_v EmptySpace v)
```

Proportions over infinite sets of infinitesimals can be well defined using random sampling:

$$[\![\mathsf{Most}\,R\,S\,]\!]_m \Leftrightarrow \lim_{K \to \infty} \mathsf{E}_{D_K \sim \pi} \frac{|D_K \cap [\![R]\!]_m \cap [\![S]\!]_m|}{|D_K \cap [\![R]\!]_m|} > 0.5$$

(Here  $\mathsf{E}_{D_K \sim \pi}$  ... is expected value of ... for K-element set  $D_K$  randomly drawn from distribution  $\pi$ .) (Think of this as setting K high enough to be reliable and ensure a non-zero denominator.) This lets us use the same quantifier functions (e.g. Most) for objects and substances.

# 2.7 Extra: propositions about arbitrary probabilities

We can further generalize quantifiers as cardinal (Count) and proportional (Ratio) quantifiers.

We can yet further generalize these as **upward-entailing**  $(Q_{\geq})$  and **downward-entailing**  $(Q_{\leq})$ :

with variants  $Q_{=}$ ,  $Q_{<}$ ,  $Q_{>}$  for other comparison operators defined in terms of these:

$$Q_{=} n R S \Leftrightarrow Q_{\leq} n R S \wedge Q_{\geq} n R S$$

$$Q_{<} n R S \Leftrightarrow Q_{\leq} n R S \wedge \neg Q_{\geq} n R S$$

$$Q_{>} n R S \Leftrightarrow Q_{\geq} n R S \wedge \neg Q_{\leq} n R S$$

Now we can express arbitrary claims about probability:

P(Edible | Nut) > .20 
$$\Leftrightarrow$$
 Ratio<sub>></sub> .20 ( $\lambda_x$  Nut  $x$ ) ( $\lambda_x$  Edible  $x$ )

# 2.8 Extra: intensions (propositions about propositions)

We now have a formal system to reason about complex ideas based on sets of entities or tuples. But what if we have to do something when someone *wants to eat*, where *to eat* is a proposition? Propositions denote truth values, but the person doesn't want 'False' (whatever that would mean). So, define argument propositions as **intensions** – sets of satisfying possible worlds [Carnap, 1947]:

(Worlds are completely specified. World models may be incomplete, subsuming many worlds.)
(Intensions are intersected with world models to ensure domains of e.g. entity constants match.)
An intension may then **entail** another if its satisfying possible worlds are a subset of the other's:

[Entail 
$$i j$$
]<sub>m</sub>  $\Leftrightarrow$   $i \subseteq j$ 

These sets would be hard to calculate! Fortunately we can define entailment in terms of structure! We reason about these, e.g. test if claim i is in some class j, by *simplifying* rather than enumerating (where  $Q \in \{\text{Count}, \text{Ratio}\}$  and  $\langle w, x, x', ... \rangle$  is a tuple of a possible world w and entities  $x, x', ... \in w$ ):

more specific intension 
$$i$$
 more general intension  $j$ 

$$\overbrace{\{\langle w,..\rangle \mid \llbracket \varphi \land \chi \rrbracket_w\}} \subseteq \overbrace{\{\langle w,..\rangle \mid \llbracket \psi \rrbracket_w\}} \qquad \text{if} \qquad \{\langle w,..\rangle \mid \llbracket \varphi \rrbracket_w\} \subseteq \{\langle w,..\rangle \mid \llbracket \psi \rrbracket_w\}$$

```
 \{\langle w, ... \rangle \mid \llbracket Q_{\geq} n R S \rrbracket_{w} \} \subseteq \{\langle w, ... \rangle \mid \llbracket Q_{\geq} n' R S \rrbracket_{w} \} \qquad \text{if} \qquad n \geq n' 
 \{\langle w, ... \rangle \mid \llbracket Q_{\leq} n R S \rrbracket_{w} \} \subseteq \{\langle w, ... \rangle \mid \llbracket Q_{\leq} n' R S \rrbracket_{w} \} \qquad \text{if} \qquad n \leq n' 
 \{\langle w, ... \rangle \mid \llbracket Q_{\geq} n R (\lambda_{x} \varphi) \rrbracket_{w} \} \subseteq \{\langle w, ... \rangle \mid \llbracket Q_{\geq} n R (\lambda_{x} \psi) \rrbracket_{w} \} \qquad \text{if} \qquad \{\langle w, ..., x \rangle \mid \llbracket \varphi \rrbracket_{w} \} \subseteq \{\langle w, ..., x \rangle \mid \llbracket \psi \rrbracket_{w} \} 
 \{\langle w, ... \rangle \mid \llbracket Q_{\leq} n R (\lambda_{x} \varphi) \rrbracket_{w} \} \subseteq \{\langle w, ... \rangle \mid \llbracket Q_{\leq} n R (\lambda_{x} \psi) \rrbracket_{w} \} \qquad \text{if} \qquad \{\langle w, ..., x \rangle \mid \llbracket \varphi \rrbracket_{w} \} \supseteq \{\langle w, ..., x \rangle \mid \llbracket \psi \rrbracket_{w} \}
```

For example:

This kind of reasoning by simplifying is sometimes called **natural logic** [van Benthem, 1986].

Entailment predicates can be used to evaluate if a desired intension *i* is in some class *j*:

```
All (\lambda_t \text{ CurrentTime } t)
(\lambda_t \text{ All } (\lambda_c \text{ Clerk } c)
(\lambda_c \text{ All } (\lambda_a \text{ Person } a)
(\lambda_a \text{ All } (\lambda_x \text{ Have } t \text{ c } x \land \text{Some } (\lambda_j \text{ IntensionOfCount}_{\geq} j \text{ 1 } (\lambda_u \text{ ConsecutiveTime } t \text{ u})
(\lambda_u \text{ Eat } u \text{ a } x))
(\lambda_j \text{ Some } (\lambda_i \text{ Want } t \text{ a } i)
(\lambda_i \text{ Entail } i \text{ } j)))
(\lambda_x \text{ Give } t \text{ c } a \text{ } x))))
```

(If a clerk has something, and someone wants it [perhaps among other things], give it to them.)

Here, even if the intension i that the agent wants contains other conjuncts, it still entails j:

```
Some (\lambda_i \text{ IntensionOfCount}_{\geq} i \ 1 \ (\lambda_u \text{ ConsecutiveTime } 10:00:00 \ u) (\lambda_u \text{ Eat } u \text{ Me Apple} 1 \land \text{Drink } u \text{ Me Juice} 1)) (\lambda_i \text{ Want } 10:00:00 \text{ Me } i)
```

So if the above is true, the clerk will recognize that I want to eat an apple and give it to me.

#### Practice: intensions

Using the non-intensional and intensional quantifier functions above and the predicates Kid k, Car c, Time t, Own t k c, and Want k i, write a lambda calculus expression stating that every kid wants to own a car at some point in time (but they don't have a particular car in mind). Note: only quantifier functions can take lambda functions ( $\lambda_x$ ...) as arguments; all the predicates can only take entity variables as arguments.

# 2.9 No need for other operators

Now we're done! Generalized quantifiers are powerful enough that we don't need other operators.

We can use a 'None' quantifier (and a uniquely satisfied predicate 'Unit') to implement negation:

```
¬ ItsRainy \Leftrightarrow None (\lambda_x \text{ Unit } x) (\lambda_x \text{ ItsRainy})
```

We can use negation to implement **disjunction** (via DeMorgan's law):

```
ItsCloudy ∨ ItsSunny ⇔ ¬ ((¬ ItsCloudy) ∧ (¬ ItsSunny))
```

And we can use disjunction to implement **implication** (via double negation law):

or more directly using an 'All' quantifier:

```
ItsRainy \rightarrow ItsCloudy \Leftrightarrow All (\lambda_x \text{ Unit } x \land \text{ ItsRainy}) (\lambda_x \text{ ItsCloudy})
```

This simplifies our natural logic entailment!

Now we have some basics of meaning, let's see how we can represent them in the brain...

#### 2.10 Review

We defined a formal system of reasoning that consists of:

- 1. predicates
- 2. conjunctions
- 3. generalized quantifiers, to model probabilistic inference, negation, disjunction, etc.
- 4. intensions, to model propositions about propositions

Coming lectures will show how to represent these things in brains and decode them from language.

# References

[Barwise & Cooper, 1981] Barwise, J. & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4.

[Carnap, 1947] Carnap, R. (1947). *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press.

[Church, 1940] Church, A. (1940). A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5(2), 56–68.

[van Benthem, 1986] van Benthem, J. (1986). Natural logic. In *Essays in Logical Semantics*. Dordrecht, the Netherlands: Kluwer.

[von Neumann & Morgenstern, 1944] von Neumann, J. & Morgenstern, O. (1944). Theory of games and economic behavior. *Science and Society*, 9(4), 366–369.