Introduction to Semantics

Lecture 2 Supplementary Material
Sets and Functions

Yasutada Sudo
UCL
y.sudo@ucl.ac.uk

OVA 2026 at Debreceni Egyetem

1 Set Theory

- A **set** is a collection of objects.
- E.g. a set containing my colleague Klaus (Abels), the number 9, Yasu's office, and nothing else is written
$$\{\,\text{Klaus}, 9, \text{Yasu's office}\,\}$$
- Sets are defined solely by their **members** (alt.: elements). If two sets are made up of the exact same members, they are said to be equivalent.
- If m is a member of a set S, we write $m \in S$. If not, we write $m \notin S$.
- The members of a set are not ordered. The same set can be represented in multiple ways, e.g.:
$$\begin{aligned} &\{\,\text{Klaus}, 9, \text{Yasu's office}\,\} \\ =&\{\,\text{Yasu's office}, \text{Klaus}, 9\,\} \\ =&\{\,9, \text{Klaus}, 9, \text{Yasu's office}, \text{Klaus}, \text{Klaus}, 9\,\} \end{aligned}$$
The last set might look like it has more members than the first two, but that is not the case; It has exactly three *distinct* members, Klaus, 9 and Yasu's office, just like the other two sets, and as such it is identical to them.
- The set with no members $\{\ \}$ is called the **empty set** and it is often denoted by a special symbol, $\varnothing$. The empty set is a set, although this symbol does not have $\{\ \}$ around it. Note that $\varnothing$ does not mean 'nothing'!
- A set can contain other sets (including the empty set), e.g. $\{\,z, \varnothing, \{\,10, 3\,\}\,\}$.
- Note that $\in$ is immediate membership. So $10 \notin \{\,z, \varnothing, \{\,10, 3\,\}\,\}$, although $\{\,10, 3\,\} \in \{\,z, \varnothing, \{\,10, 3\,\}\,\}$.
- The size of the set A is called the **cardinality** of A and written $|A|$.[1]
e.g. $|\,\{\,3, a, \varnothing\,\}\,| = 3, \quad |\,\{\,7, \varnothing, \text{Jesse}, \{\,a, 1\,\}, \text{Berlin}\,\}\,| = 5$

[1] The idea of cardinality becomes more interesting with sets with infinite members. For example, perhaps counterintuitively, some infinite sets have greater cardinalities than other infinite sets. If you are interested in this, I recommend Chapter 4 of Partee et al. (1990) mentioned below, and the part about Gregor Cantor of the BBC documentary *Dangerous Knowledge*.

1.1 Two ways of representing sets

1.1.1 Extensional notation (enumeration)

- Simply enclose the members between '{' and '}' and put a ',' between every pair, e.g. $\{ d, 1, \varnothing, \text{Bill} \}$.
- As mentioned above, the order among the members does not matter.
- Sets with infinite members cannot be represented this way (although if there is a predictable pattern, we sometimes informally write $\{ 1, 2, 3, \ldots \}$ for the set of all natural numbers, for example).

1.1.2 Intensional notation (abstraction)

- Use a variable and state the membership condition in words, e.g. $\{ x \mid x \text{ is an even number} \}$ ($\approx$ the set of x's such that x is an even number).
- Sometimes ':' is used in place of '|', as in $\{ x : x \text{ is an even number} \}$.
- The variable name does not matter, e.g.

$$\{ x \mid x \text{ is a natural number} \} = \{ y \mid y \text{ is a natural number} \}$$

- Both finite and infinite sets can be represented this way. The same set as the very first set we encountered above can be written as $\{ z \mid z \text{ is Klaus or } 9 \text{ or Yasu's office} \}$. Note the use of 'or' here; 'and' won't give you the same set (Nothing is both Klaus and 9, for example). Note also that this set does not contain z, which is a variable here.

1.2 Relations among sets

- If every member of A is also a member of B, we say A is a **subset** of B and write $A \subseteq B$. Notice for any set A, $A \subseteq A$, because every member of A is obviously a member of A, and we consider this to apply to $\varnothing$ as well, so $\varnothing \subseteq \varnothing$.
- $\varnothing$ is a subset of every set, including $\varnothing$ itself.
- A and B are equivalent (written $A = B$) iff $A \subseteq B$ and $B \subseteq A$.
- If $A \subseteq B$ and $A \neq B$, we say A is a **proper subset** of B and write $A \subset B$.[2]
- The set of all subsets of a set A is called its **power set**, $\wp(A)$ (or sometimes 'Pow(A)', $\mathfrak{P}(A)$, etc.), e.g. $\wp(\{ a, b \}) = \{ \varnothing, \{ a \}, \{ b \}, \{ a, b \} \}$. Because $\varnothing$ is a subset of every set, $\wp(S)$ always contains $\varnothing$, whatever set S is.

1.3 Operations on sets

For any two sets A and B:

- **Intersection:** $A \cap B \coloneqq \{ x \mid x \in A \text{ and } x \in B \}$
- **Union:** $A \cup B \coloneqq \{ x \mid x \in A \text{ or } x \in B \}$

[2] Notational caveat: Confusingly, some old work uses $\subset$ for 'subset' and $\subsetneq$ for 'proper subset'. We will consistently use the more modern notation.

- **Set subtraction/Relative complementation:** $A - B = \{ x \mid x \in A \text{ and } x \notin B \}$

(the symbol '$\coloneqq$' means 'is defined to be equal to', and is often used to state definitions)

1.4 Ordered pairs and n-tuples

- Unlike members of sets, members of **ordered pairs** $\langle x, y \rangle$ are ordered. If $x \neq y$, then $\langle x, y \rangle \neq \langle y, x \rangle$. If $x = y$, then $\langle x, y \rangle = \langle y, x \rangle$.
- Ordered pairs can be defined as sets of a particular kind. There are several ways of doing so, but the most common way is Kazimierz Kuratowski's definition: $\langle x, y \rangle \coloneqq \{ \{ x \}, \{ x, y \} \}$.
- We can generalise ordered pairs to sequences of length n for $n > 1$, which are called **n-tuples**, as follows:
  - $\langle x, y, z \rangle \coloneqq \langle x, \langle y, z \rangle \rangle$
  - $\langle w, x, y, z \rangle \coloneqq \langle w, \langle x, y, z \rangle \rangle = \langle w, \langle x, \langle y, z \rangle \rangle \rangle$
  - More generally, $\langle x_1, x_2, \ldots, x_{n-1}, x_n \rangle = \langle x_1, \langle x_2, \langle \cdots, \langle x_{n-1}, x_n \rangle \cdots \rangle \rangle$
- We could alternatively use left association, as in $\langle x, y, z \rangle \coloneqq \langle \langle x, y \rangle, z \rangle$. That would make no practical difference, as long as we are consistent.

2 Functions

Functions play a particularly important role in formal semantics. Mathematically, functions are simply 'mappings' of a particular kind. You can think of a function as a machine or computer program that takes some input and returns some output. One important property of functions is that a function always maps a given input to the same output, every time that input is given to the function. In other words, functions are deterministic mappings.

As a matter of convention (due to Leonhard Euler), we denote the output of a function F when **applied to** some input (or **argument**) A by '$F(A)$'. You must be familiar with this notation from high school mathematics, e.g. if your function is $g(x) = x^2 + 2 \times x$, then $g(2) = 2^2 + 2 \times 2 = 8$. This means g is a function that maps 2 to 8, or equivalently, the function g outputs 8 when applied to 2.

In order for a mapping F to count as a function, the following must be true: for any argument A, the output $F(A)$ is unique. Here are some examples:

- R maps each city in the UK to the number of its (registered) residents (as of a particular time, say at 17:00 on 21 January 2022). This is a function because for each British city c, there's a unique number $R(c)$.
- S maps any positive or negative integer n (including 0) to n^2. This mapping is a function, because for each input value, there is only one output value. Notice that in this case there are some outputs that can be produced by different inputs, e.g. $S(2) = S(-2) = 4$. But this is fine, because what matters is the uniqueness of the output for each input.
- Z maps every number to 0. This is a function, because when an input is given, we can uniquely determine the output. The output is actually the same, no matter what the input is, but that's fine.

- G maps any person in my department to one of the languages they speak. This is *not* a function, because for some inputs, there are multiple outputs. For instance, Nausicaa Pouscoulous speaks several languages.

In order to define a function, you need to specify the following three things:

- **Domain**: the set of objects the function takes as inputs.
- **Range**: the set of objects the function returns as outputs.
- What the function does to each input.

In the case of the squaring function S above:

- The domain of S (written dom(S)) is the set of all integers (= whole numbers). This set is often denoted by $\mathbb{Z}$.
- The range of S (written ran(S)) is the set of all natural numbers (including 0), $\mathbb{N}$. (You can also say the range is $\mathbb{Z}$ in this case)
- For any integer n, $S(n) = n^2$.

The domain and range are sometimes written together as '$S : \mathbb{Z} \to \mathbb{N}$'. This means '$S$ is a function that maps any integer to a natural number'.

We can describe a function using a natural language like English, as in (1).

(1)
a. The domain of the function M is the set of British cities. M maps any British city to its mayor.
b. the function that is only defined for people and maps any person x to 1 iff x is a smoker
c. the function which maps any American politician to Donald Trump and is undefined for anything that is not an American politician
d. the function over natural numbers that maps any natural number n to $n + 5$
e. The domain of the function is $\{ \text{a}, \text{e}, \text{i}, \text{o}, \text{u} \}$ and it maps a to e, e to o, i to a, o to u, and u to i.

Alternatively, we can give a more graphic represent of a function. For example, (1e) can be written as follows.

$$\begin{bmatrix} \text{a} & \to & \text{e} \\ \text{e} & \to & \text{o} \\ \text{i} & \to & \text{a} \\ \text{o} & \to & \text{u} \\ \text{u} & \to & \text{i} \end{bmatrix}$$

This representation is not very convenient when the domain of the function is infinite. For instance, the best we can do for (1d) is something like:

$$\begin{bmatrix} 1 & \to & 6 \\ 2 & \to & 7 \\ 3 & \to & 8 \\ 4 & \to & 9 \\ & \vdots & \end{bmatrix}$$

In this case there's a clear pattern, so one can guess what'll happen to other natural numbers, but in the general case there is no guarantee that there's a pattern that can be represented this way. For example, one can think of a function that maps any natural number n to n times the nth decimal digit of π. That cannot be represented in the same way.

Another way of representing functions is to understand a function as a set of ordered pairs. In terms of ordered pairs, (1e) will look as follows.

$$\{\ \langle \mathrm{a}, \mathrm{e} \rangle,\ \ \langle \mathrm{e}, \mathrm{o} \rangle,\ \ \langle \mathrm{i}, \mathrm{a} \rangle,\ \ \langle \mathrm{o}, \mathrm{u} \rangle,\ \ \langle \mathrm{u}, \mathrm{i} \rangle\ \}$$

Since sets with infinite members can be defined using the abstraction notation, we can represent the function in (1d) as follows.

$$\{ \langle x, y \rangle \mid x \text{ is a natural number and } y = x + 5 \}$$

To make your life more complicated, semanticists often use so-called λ-notation ('λ' is the Greek letter lambda). This will be introduced in Lecture 2.