Reasoning is the process of making inferences (drawing conclusions) from some information.

Three kinds of reasoning can be distinguished:

- **Inductive reasoning**: generalizing from a set of observations to a rule (e.g., observe a number of white swans, conclude *all swans are white*);
- **Deductive reasoning**: draw conclusions from premises using logical rules (e.g., given *Aristotle is a man* and *all man are mortal* conclude *Aristotle is mortal*);
- **Abductive reasoning**: reason from a conclusion (or effect) to an explanation (or cause) (e.g., a mediocre athlete performs exceptionally, conclude he is doped).

Crucially, only *deductive reasoning* guarantees that the conclusion is correct if the premises are correct, based on the rules of logic.

- Inductive reasoning: counterexamples may exists that render the conclusion invalid (e.g., there is a black swan, but we haven’t observed it);
- Abductive reasoning: no guarantee that the inferred explanation is correct (there could be other explanations).

This lecture will focus on deductive reasoning, and on reasoning with syllogisms in particular.
In the psychological literature, deductive reasoning has been studied using the following types of problems:

**Transitive reasoning** problems involve transitive relations. Example: *A is taller than B and C is shorter than B*. What follows about *A* and *C*?

**Conditional reasoning** problems involve conditional statements. Example: *if it is dark then the street lights will be on and the street lights are on*. What follows?

**Syllogistic reasoning** problems involve statements about categories. Example: *all lions are savage animals and all lions are cats*. What follows about the relation between *savage animals* and *cats*?

Syllogisms are inferences from two *premises* to a *conclusion*:

- premises and conclusion are expressed as set-theoretic relationships between categories;
- the conclusion does not follow from experience, but just from the structure of the set-theoretic relations in the premises.

Example:

- Some artists are beekeepers
- No beekeepers are chemists
- Some artists are not chemists

This follows because the artists who are beekeepers cannot be chemists.

Syllogisms have a fixed structures:

- each premise must have one of four *quantifiers*: all, no, some, some . . . not. Quantifiers express set-theoretic relationships;
- the quantifiers relate two *terms*, one of which, the *middle term* appears in both premises. Terms express categories;
- the conclusion also contains one of four quantifies and relates the remaining two terms, the *end terms*.

Syllogisms come in four *figures*, depending on where in the premises the middle term comes in relation to the end terms.

Example: figure *ab/bc*: middle term comes second in first premise and first in second premise (see beekeeper example above). Figures *ba/bc, ba/cb, ab/cb* also possible.
Figural effect: experiments show a bias towards conclusions whose end-term order is related to the figure of the premises:

- some A are B
- all B are C
- some A are C

preferred

- some A are B
- all B are C
- some C are A
dispreferred

In the ab/bc figure, the bias is towards ac conclusions, in the ba/cb figure it is towards ca conclusions, and in the other figures the numbers of ac and ca conclusions are equal.

An influential theory of deductive reasoning is Johnson-Laird and Byrne’s (1991) Mental Models theory:

- assumes that people create a mental model (an arrangement of symbols) to determine which conclusions follow from premises;
- memory limitations and strategic biases explain why certain conclusions are easier to draw than others;
- has been applied to syllogisms in particular;
- can explain effects such as the figure effect.

We will look at modeling syllogistic reasoning using Mental Models.

Model Construction

General approach: build a tabular representation of the situation described by the premises, then revise this representation, and read off the conclusions.

Example:

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We represent the first premise some A are B as:

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each row represents an individual which is both an A and a B (the number of individuals is unimportant). Other individuals are also possible (hence the dots).

Now augment the model with the second premise all B are C:

We revise the model by adding other possible individuals:

As some A are B, it’s possible to have As that are not Bs, and Bs that are not As. As All B are C, the Bs that are not As must be Cs, but also Cs that are not Bs are possible.

The only conclusions that holds in both models (original and revised) is some A are C and some C are A.
Predictions

Key assumption of mental models theory: models are constructed in a buffer with \textit{first-in first-out} (FIFO) access.

This explains the figural effect:

- the premise with the end term in subject position is entered into the model first;
- for \textit{ab} and \textit{bc} syllogisms, the \textit{a} term is entered into the model before the \textit{c}, so FIFO access results in preference for the \textit{ac} conclusion;
- for \textit{ba} and \textit{cb} syllogisms, the \textit{c} term is entered before the \textit{a} term, resulting in a preference for the \textit{ca} conclusion;
- for \textit{ab} and \textit{cb} syllogisms, the either both or none of the end terms is in subject position, so there is no preference.

Building a Mental Model

Let’s sketch an implementation of the mental models theory in Cogent. Cooper (2002: Ch. 5) assumes the following architecture:

- \textit{Problem Buffer} for input of premises and output of conclusions;
- \textit{Scheduler} controls task sequence: build model of each premise, draw conclusions, revise model, again draw conclusions;
- \textit{Mental Model} buffer contains the model;
- \textit{Build Initial Model} process constructs the model and \textit{Draw Conclusions} process generates conclusions.

We won’t cover \textit{Revise Model} which uses of the \textit{Annotation} buffer to revise initial model based on counterexamples.

Scheduler

The \textit{Scheduler} process first initialises the model:

\begin{verbatim}
IF not initialised(_,_) is in Problem Buffer
premise(Premise1) is in Problem Buffer
premise(Premise2) is in Problem Buffer
Premise1 is distinct from Premise2
extract_term(Premise1,Premise2,initial,X)
THEN send initialise(X) to Build Initial Model
add initialised(Premise1,Premise2) to Problem Buffer
\end{verbatim}

and then integrates the next premise into the model:

\begin{verbatim}
IF initialised(Premise1,Premise2) is in Problem Buffer
extract_integration_order(Premise1,Premise2,Order)
promise_to_integrate(Order,Premise)
extract_direction(Premise1,Premise2,Order,Direction)
THEN send premise(Premise,Direction) to Build Initial Model
add integrated(Premise) to Problem Buffer
\end{verbatim}
Then it triggers conclusion drawing:

**IF** initialised(Premise1, Premise2) is in **Problem Buffer**
**integrated(Premise1)** is in **Problem Buffer**
**extract_term(Premise1, Premise2, middle, Middle)**

**THEN** send initial_concs(Middle) to **Build Initial Model**

This relies on conditions of the following type:

- **extract_term**([Q1,A,B],[Q2,B,C],initial,A).
- **extract_integration_order**([Q1,A,B],[Q2,B,C],
  [[Q1,A,B],[Q2,B,C]],
  **extract_direction**([Q1,A,B],[Q2,B,C],[Q1,A,B],forward).

---

Then individuals (based on no, some, some ... not quantifiers) are added:

**TRIGGER** premise([no,X,Y],forward)
**IF** data(.,X,X) is in **Mental Model**
NewInd is a new symbol with base ’I’
**THEN** add data(NewInd,Y,Y) to **Mental Model**
add exhaust(X,Y) to **Annotations**
add exhaust(Y,X) to **Annotations**

The **Annotations** buffer keeps track of exhaustive predicates, i.e., predicates that apply to all individuals.

---

The **Build Initial Model** process adds the term to **Mental Model**
(a table buffer) as data(Row, Column, Term):

**TRIGGER** initialise(Term)
**IF** Ind1 is a new symbol with base ’I’
Ind2 is a new symbol with base ’I’
**THEN** add data(Ind1, Term, Term) to **Mental Model**
add data(Ind2, Term, Term) to **Mental Model**

Then exhaustive links (based on all quantifiers) are added:

**TRIGGER** premise([[all,X,Y],forward])
**IF** data(Ind, X, X) is in **Mental Model**
**THEN** add data(Ind, Y, Y) to **Mental Model**
add exhaust(X,Y) to **Annotations**

---

The **Draw Conclusions** process draws conclusions in FIFO order:

**TRIGGER** initial_concs(B)
**IF** generate_conclusions(B, Conc)
  truth_condition(Conc)
**THEN** add conclusion(Conc) to **Problem Buffer**
send initial_concs(B) to **Draw Conclusions**

The following predicate generates all possible conclusions so that their truth conditions can be tested:

```
generate_conclusion(Mid, [Quant, Subj, Pred]) :-
  get_end_term(Mid, Subj),
  get_end_term(Mid, Pred),
  Subj is distinct from Pred
  Quant is a member of [all, no, some, some not]
```
Conclusions

Now we need predicates that check the truth of a conclusion against the mental model:

\[
\text{truth\_condition([[all, S, P]]) :-
not data(Ind, _, S) is in Mental Model}
not data(Ind, _, P) is in Mental Model
\]

\[
\text{truth\_condition([[no, S, P]]) :-
not data(Ind, _, S) is in Mental Model}
data(Ind, _, P) is in Mental Model
\]

\[
\text{truth\_condition([[some, S, P]]) :-
exists data(Ind, _, S) is in Mental Model}
data(Ind, _, P) is in Mental Model
\]

\[
\text{truth\_condition([[somenot, S, P]]) :-
exists data(Ind, _, S) is in Mental Model}
not data(Ind, _, P) is in Mental Model
\]

To ensure that it draws all and only valid conclusions, the implementation must revise the initial model, retest the conclusions, and delete ones for which there are counterexamples.

Alternative to mental models: reasoning based on Euler circles:

- uses diagrammatic (instead of propositional) representation of the model;
- no revision required; constructs single model that integrates both premises;
- only applied to syllogistic reasoning (while mental models have been more generally applied).

Discussed in Cooper (2002: Ch. 5) in more detail.

Summary

- Types of reasoning: inductive, deductive, abductive;
- syllogistic reasoning is a form of deductive reasoning;
- that some syllogisms are easy, others are difficult;
- figure effect: depending on the figure of the syllogism (its sequence), some conclusions are preferred over others;
- mental models theory: people solve syllogisms by creating a model that represents the individuals in the premises, and then draw conclusions from that;
- can explain the figure effect in terms of memory access;
- model revision is required in order to deal with counterexamples.

References
