# COMP9414: 人工智能

## 第4a讲：知识表示

韦恩-沃布克

电邮：w．wobcke@unsw．edu．au





## 本讲座

- 知识表示和逻辑
- 逻辑上的争论
- 命题逻辑
  - 语法
  - 语义学
- 有效性、等价性、可满足性、实体性





## 知识水平

知识水平假说。存在一个独特的计算机系统层次，紧邻符号层次，其特点是以知识为媒介，以理性原则为行为规律。

理性原则。如果一个代理人知道它的一个行动将导致它的一个目标，那么这个代理人将选择这个行动。

知识。任何可以归属于一个代理人的东西，使其行为可以根据理性原则进行计算。

"知识水平"（Newell，1982年）。





## 知识表征

- **自然演**绎法的推理

- **任何代理人都可以在不同**层面上进行描述
  - 知识水平（归属于代理人的知识）。
  - 逻辑层面（操作知识的算法）。
  - 实施层面（如何实施算法）。
- 知识表征关注的是以计算机可操作的方式明确表达知识（供代理人在推理中使用）--与Newell的观点不同。
- 推理试图利用这些知识并进行推理（例如，回答查询，确定从知识中得出的事实，决定做什么， 等等）--作为代理架构的一部分。




新南威尔士大学
©W.Wobcke等人，2019-2022年

## 知识表示和推理

- 一个基于知识的代理，其核心是一个知识库
- 知识库是一个明确的关于某些领域的句子集，用合适的形式化表示语言表达。
  - 句子表达事实（真实）或非事实（虚假）。
  - 所以 "知识库 "最好称为 "信念库"
- 基本问题
  - 我们如何写下关于某个领域/问题的知识？
  - 我们如何实现推理自动化，以推导出新的事实或确保知识库的一致性？





## 激励性的例子--本体论

阿帕克本体论

- 阿什拉夫-加尼就是加尼总统--平等
- 阿什拉夫-加尼是阿富汗的总统 - 角色
- 阿什拉夫-加尼在政府中--属于
- 楠格哈尔省是一个省，是一种
- 楠格哈尔省位于阿富汗境内--属于
- 轰炸意味着攻击--语言学意义/语义学





## 为什么是正式的语言--而不是英语？

- 自然语言表现出模糊性
  "渔夫去了银行"(词条)
  "男孩用望远镜看到一个女孩"(结构)
  "桌子无法通过门口，因为它太[宽/窄]了"（共同参照）。
- 含糊不清使人难以解释短语/句子的含义
  - 但也使推理更难定义和计算
- 符号逻辑是一种语法上无歧义的语言（最初是为了使数学推理正规化而开发的）。





## 语法与语义

本体论=此类事实的集合

语法
描述知识表示语言中的合法句子（例如，在算术表达式的语言中$x<4$）。

语义学指的是句子的意义。将句子（和句子片段）与句子所涉及的世界的各个方面联系起来。语义学指的是句子与 "真实世界"或世界的某些模型的关系。句子的语义属性包括真与假（例如，$x<4$在$x=3$时为真，$x=5$时为假）。名称和描述的语义属性包括指称物。

注意：一个句子的含义不是该句子的内在含义。需要有一个解释来确定句子的含义。解释是在一个语言社区之间达成的。

# 主张

- 命题是可以是真的或假的实体（事实或非事实）。
- 使用普通的陈述句来表达（不是问题）。
  - "天空是蓝色的
    "表达的命题是：天空是蓝色的（此时此地）。这个命题是真的吗？
- 实例
  "苏格拉底是个秃子"（假设 "苏格拉底"、"秃子
  "都有明确的定义）"汽车是红色的"（要求 "汽车
  "被识别出来）
  "苏格拉底是秃头，汽车是红色"（复杂命题）。
- 在命题逻辑中，用单个字母表示命题，是一种缩写方案，例如，$P$：苏格拉底是秃子
- **重要的是。推理是独立于命题子结构的!**





# 逻辑上的争论

一个论证将一组前提与一个结论联系起来

– 如果结论必然来自于前提，则有效

所有人类都有两只眼睛
简是一个人
所以简有两只眼睛

所有的人都有4只眼睛





# 逻辑 论点

一个论证将一组前提与一个结论联系起来

- 当前提都是真的时候，如果结论可能是假的，那就是无效的

所有的人都有两只眼睛
简有两只眼睛
所以简是人类

没有人有4只眼睛
简有2只眼睛
因此，简不是人

- 两者都是（逻辑上）不正确的无效论点
- **哪些说法是真的/假的？**





# 命题逻辑

简是一个人 所以简有4只眼睛

- 两者都是（逻辑上）正确的有效论点
- **哪些说法是真的/假的？**

- 用字母代表 "基本 "命题；用 "不是"、"和"、"或"、"意味着"、"iff "等运算符将它们组合成更复杂的句子

- 命题连接词。

| | | | |
|---|---|---|---|
| $\neg$ | 否定 | $\neg P$ | "不是P" |
| $\wedge$ | 合并 | $P \wedge Q$ | "P和Q" |
| $\vee$ | 分歧点 | $P \vee Q$ | "P或Q" |
| $\rightarrow$ | 意味着 | $P \rightarrow Q$ | "如果P，那么Q" |
| $\leftrightarrow$ | 双重性 | $P \leftrightarrow Q$ | "P当且仅当Q" |



新南威尔士大学
©W.Wobcke等人，2019-2022年

## 从英语到命题逻辑

- "天是蓝的"，这不是事实。$\neg B$
  (或者说 "天空不蓝")
- "天是蓝的，草是绿的"。$B \wedge G$
- "要么天是蓝的，要么草是绿的"。$B \vee G$
- "如果天空是蓝色的，那么草就不是绿色的"。$B \rightarrow \neg G$
- "天空是蓝色的，当且仅当草是绿色的"。$B \leftrightarrow G$
- "如果天空是蓝色的，那么如果草不绿，植物就不会生长"。$B \rightarrow (\neg G \rightarrow \neg P)$





## 提高可读性

- $(P \rightarrow (Q \rightarrow (\neg(R))))$ vs $P \rightarrow (Q \rightarrow \neg R)$
- 省略括号的规则
  - 尽可能地省略括号（也许下面的最后一个例子除外！）。
  - 顺序从高到低是。$\neg, \wedge, \vee, \rightarrow, \leftrightarrow$
  - 所有二元运算符都是左联的（所以$P \rightarrow Q \rightarrow R$缩写为 $(P \rightarrow Q) \rightarrow R$）
- 问题
  - $(P \vee Q) \vee R$ (总是) 与 $P \vee (Q \vee R)$ 相同吗？
  - $(P \rightarrow Q) \rightarrow R$（总是）与$P \rightarrow (Q \rightarrow R)$ 相同吗？





## 真值表语义

- 连接词的语义可由真值表给出

| $P$ | $Q$ | $\neg P$ | $P \wedge Q$ | $P \vee Q$ | $P \rightarrow Q$ | $P \leftrightarrow Q$ |
|---|---|---|---|---|---|---|
| 真 | 真 | 假的 | 真 | 真 | 真 | 真 |
| 真 | 假的 | 假的 | 假的 | 真 | 假的 | 假的 |
| 假的 | 真 | 真 | 假的 | 真 | 真 | 假的 |
| 假的 | 假的 | 真 | 假的 | 假的 | 真 | 真 |

- **对变量的 "真/假 "的每一种可能分配都有一行。**
- 重要：$P$和$Q$是任何句子，包括复杂的句子。





## 例子 - 复杂句子

| $R$ | $S$ | $\neg R$ | $R \wedge S$ | $\neg R \vee S$ | $(r \wedge s) \rightarrow (\neg r \vee s)$ |
|---|---|---|---|---|---|
| 真 | 真 | 假的 | 真 | 真 | 真 |
| 真 | 假的 | 假的 | 假的 | 假的 | 真 |
| 假的 | 真 | 真 | 假的 | 真 | 真 |
| 假的 | 假的 | 真 | 假的 | 真 | 真 |

因此，$(R \wedge S) \rightarrow (\neg R \vee S)$ 是一个同义反复。

## 定义

- **如果一个句子在所有可能的 "真"/"假** "变量分配下都是"真"，则该句子是有效的（例如，$P \lor \neg P$）。
- 同义词是一个有效的句子
- 如果两个句子有相同的真值表，那么它们就是等价的，例如：。$P \land Q$和$Q \land P$
  - *所以*当且仅当$P \leftrightarrow Q$有效时，$P$与$Q$是等价的。
- 如果存在一些对其变量的真/假分配，且该句子为真，则该句子是可满足的。
- 如果一个句子是不可满足的（例如，$P \land \neg P$），那么它就是不可满足的。
  - 句子对其变量的所有真/假分配都是假的
  - *所以*，当且仅当¬P是不可满足的时候，$P$是一个同义词。





## 物质影响

- 只有当$P$为真，$Q$为假时，$P \to Q$才会评估为假。
- $P \to Q$等同于$\neg P \lor Q$：实质暗示
- 英语的用法常常表明前因后果之间的联系。($P$)和结果($Q$)--这并没有反映在真值表中
- 实例
  - $(P \land Q) \to Q$是一个同义词，适用于任何$Q$
  - $P \to (P \lor Q)$对任何$Q$*都*是同义词。
  - $(P \land \neg P) \to Q$对任何$Q$来说都是同义词。





## 逻辑等价物--全部有效

| | | |
|---|---|---|
| 交换性。 | $p \land q \leftrightarrow q \land p$ | $p \lor q \leftrightarrow q \lor p$ |
| 关联性。 | $p \land (q \land r) \leftrightarrow (p \land q) \land r$ | $p \lor (q \lor r) \leftrightarrow (p \lor q) \lor r$ |
| 分配性。 | $p \land (q \lor r) \leftrightarrow (p \land q) \lor (p \land r)$ | $p \lor (q \land r) \leftrightarrow (p \lor q) \land (p \lor r)$ |
| 寓意。 | $(p \to q) \leftrightarrow (\neg p \lor q)$ | |
| 无效的。 | $p \land p \leftrightarrow p$ | $p \lor p \leftrightarrow p$ |
| 双重否定。 | $\neg\neg p \leftrightarrow p$ | |
| 矛盾。 | $p \land \neg p \leftrightarrow$ FALSE | |
| 不包括中间。 | | $p \lor \neg p \leftrightarrow$ TRUE |
| 德-摩根。 | $\neg(p \land q) \leftrightarrow (\neg p \lor \neg q)$ | $\neg(p \lor q) \leftrightarrow (\neg p \land \neg q)$ |





## 等效性的证明

让$P \Leftrightarrow Q$表示 "$P$等同于$Q$" ($P \Leftrightarrow Q$不是一个公式) 那么$P \land (Q \to R) \Leftrightarrow \neg(P \to Q) \lor (P \land R)$

| | | | |
|---|---|---|---|
| $P \land (Q \to R)$ | $\Leftrightarrow$ | $P \land (\neg Q \lor R)$ | [含义] |
| | $\Leftrightarrow$ | $(p \land \neg q) \lor (p \land r)$ | [分配性] |
| | $\Leftrightarrow$ | $(\neg\neg p \land \neg q) \lor (p \land r)$ | [双重否定] |
| | $\Leftrightarrow$ | $\neg(\neg p \lor q) \lor (p \land r)$ | [德-摩根] |
| | $\Leftrightarrow$ | $\neg(p \to q) \lor (p \land r)$ | [含义] |

假设替代：如果$A \Leftrightarrow B$，则在任何子公式中用$B$替换$A$

假设等价关系是传递性的：如果$A \Leftrightarrow B$和$B \Leftrightarrow C$，则$A \Leftrightarrow C$





新南威尔士大学

©W.Wobcke等人，2019-2022年

# 诱惑

- 如果$S$中的所有公式都是 "真"， $P$也是 "真"， 那么$S$就包含了$P$ ($S$ |= $P$)。
  - 语义学定义--涉及真理（非证明）。
- 通过计算$S$和$P$的真值表来计算$S$ |= $P$
  - 句法概念 - 涉及计算/验证
  - 并非总是如此容易计算（这是多低的效率？）
- 同义反复是必然性的一个特例，其中$S$是空集。
  - 真值表的所有行都是真





# 简单的要求

将$P$ |= $Q$写为{P}|= $Q$

| | |
|---|---|
| $p \wedge q$ \|= | $pp \wedge q$ \|= $q$ |
| $p$ \|= $p \vee$ | $qq$ \|= $p \vee q$ |
| $P$ \|=¬¬P | ¬¬ÃÃÃ |

{P, $P \rightarrow Q$}|= $Q$

如果$P$ |= $Q$， 那么 |= $P \rightarrow Q$





# 实例

| $P$ | $Q$ | $P \rightarrow Q$ | $Q$ |
|---|---|---|---|
| 真 | 真 | 真 | 真 |
| 真 | 假的 | 假的 | 假的 |
| 假的 | 真 | 真 | 真 |
| 假的 | 假的 | 真 | 假的 |

因此{P, $P \rightarrow Q$}|= $Q$

- 在唯一一行$P$和$P \rightarrow Q$都是真（第1行）的地方，$Q$也是真（这里$S$是集合{P， $P \rightarrow Q$}）。

注：$P \rightarrow Q$这一栏是用真值表的定义从$P$和$Q$中计算出来的，而$Q$又被用来检查必然性





# 实质性问题 - 句法

| $R$ | $S$ | ¬R | $R \wedge S$ | ¬R $\vee S$ | $(r \wedge s) \rightarrow (\neg r \vee s)$ |
|---|---|---|---|---|---|
| 真 | 真 | 假的 | 真 | 真 | 真 |
| 真 | 假的 | 假的 | 假的 | 假的 | 真 |
| 假的 | 真 | 真 | 假的 | 真 | 真 |
| 假的 | 假的 | 真 | 假的 | 真 | 真 |

因此 |= $(R \wedge S) \rightarrow (\neg R \vee S)$

# 模型

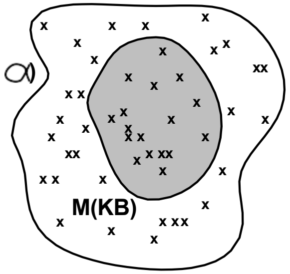
也可以用模型来思考，正式的结构化的解释，根据这些解释来评价真理。

- 对于命题逻辑，一个模型是真值表的一行，如果α在$M$

中为真，那么模型$M$就是一个句子α的模型。

让$M$(α)为α的所有模型的集合

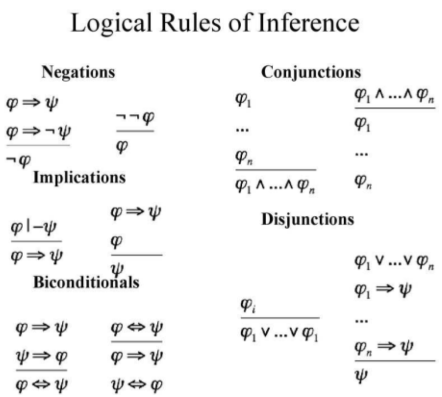
那么，当且仅当$M$（$KB$）⊆$M$（α） M(








# 自然演绎证明








# 自然扣除法实例

$$\dfrac{\dfrac{\dfrac{\dfrac{\overline{A\to(B\to C)}^{\,2}\qquad \dfrac{\overline{A\wedge B}^{\,1}}{A}}{B\to C}\qquad \dfrac{\overline{A\wedge B}^{\,1}}{B}}{C}}{A\wedge B\to C}\,1}{(A\to(B\to C))\to(A\wedge B\to C)}\,2$$





# 总结

注释：⊢表示证明；⇒是我们的→。

- 用形式语言避免了自然语言的模糊性
- 使得（保真）必然性的形式化成为可能
- 命题逻辑。最简单的真与假的逻辑
- 基于知识的系统。一阶逻辑
- 自动推理。如何计算连带关系（推理）？
- 本课程中没有学习的许多逻辑学




新南威尔士大学
©W.Wobcke等人，2019-2022年