更新语义中的常识[1]

弗兰克·斐尔曼

摘要：本文的目标为以下两方面：（i）引入更新语义的框架并解释哪种现象能在其中成功的被分析；（ii）给出其中一种现象——常识推理——的详细分析。

1．介绍：更新语义的框架

逻辑有效性的标准定义是这样的：一个论证（argument）是有效的，如果当其结论不真时，其前提（可能多个）不能都真。许多逻辑理论都是以这一有效性定义为起点发展起来的。随之而来的，这些理论的核心包含一个这样的真条件的规范。本文中理论的核心不包含这一真条件规范标语："如果你知道一个句子在什么条件下真，那么你就知道一个句子的含义"被"如果你知道该句对任何接受该句所传达的信息的主体的信息状态所带来的改变时，你知道一个句子的含义"所替代。[1]这样，含义变成了一个动态的概念：一个句子的含义是一个信息状态上的算子。

为了给语言$\mathscr{L}$定义一个更新语义，我们需要一个相关信息状态集Σ，函数[ ]：给每个句子ϕ指派一个Σ上的算子$[\varphi]$。三元组$<\mathscr{L}, \Sigma, [\ ]>$称为更新系统。如果$\sigma$是一个状态，$\varphi$是一个句子，我们用"$\sigma[\varphi]$"表示用$\varphi$更新$\sigma$的结果。因为$[\varphi]$是函数，$\sigma$是论证，写成$[\varphi](\sigma)$似乎更自然，但后缀式更方便文字上的处理。现在，我们可以把用句子序列$\psi_1,\ldots, \psi_n$更新σ记为"$\sigma[\psi_1]\ldots[\psi_n]$"。

一个重要的概念是接受（acceptance）。设σ是任意状态，φ是任意句子。考虑状态$\sigma[\varphi]$。这一状态在大多数情况下与σ都不相同，但随时可能出现$\sigma[\varphi]=\sigma$，如果这样，由φ所传达的信息已经由σ总结出来了。这种情况下，我们记为$\sigma\Vdash\varphi$，并说φ在σ中被接受了。

1.1 并非总是成立的限制

"更新语义"这个词可能会有些误导，似乎暗示着当你用φ来更新你的信息状态时，所有你不得不做的就是把φ的信息内容添加到你已有的信息之上。

定义 1.1 一个更新系统$<\mathscr{L}, \Sigma, [\ ]>$是累积的，当且仅当，$\Sigma$中有状态 **0**（最小状态）以及$\Sigma$上的一个二元运算+，使得

(i) 运算+有如下性质：

$$\mathbf{0}+\sigma=\sigma$$

$$\sigma+\sigma=\sigma$$

$$\sigma+\tau=\tau+\sigma$$

本文作者 Frank Veltman，为荷兰阿姆斯特丹大学逻辑、语言与计算研究所教授，知名逻辑学家。原文：Defaults in Update Semantics, Journal of philosophical logic, Springer, volume 25, Number 3, 221-261, 1996。
译者：张立英，哲学博士，中央财经大学 现代逻辑研究所 副教授。

[1] 译者注：default 在本文中被翻译为常识。

$$(\rho+\sigma)+\tau=\rho+(\sigma+\tau)$$

(ii) 对每一个句子φ和状态σ，$\sigma[\varphi]=\sigma+\mathbf{0}[\varphi]$

(i)成立时Σ被称为信息格。如果$\sigma+\tau=\tau$，我们记为$\sigma\leqslant\tau$，称为τ不弱于σ。

只要人们处理的是可以用经典更新系统所刻画的现象，动态的研究不会比静态的研究提供更多的东西。这些案例下，人们可以把$\mathcal{L}$中的每个句子φ与一个静态的含义联系起来——$\mathbf{0}[\varphi]$，表示φ的信息内容——从而以这种形式来定义φ的动态含义。

如果要一个更新系统是*累积*的，还需满足不同的限制条件。其中一个是：对每个σ，$\sigma[\varphi]$都应被定义。本文中所讨论的系统都有这一性质，但不难想象，有不满足这一条件的现象，以下的例子可作为参考。

“他只是在开玩笑”

如果不清楚说话人指的是谁，则听者不会知道对这一陈述做出何种反应。或者考虑预设。更新语义的框架这一概念给出了一个自然的解释：

φ预设ψ当且仅当对每个状态σ，$\sigma[\varphi]$是已定义的仅当$\sigma\Vdash\varphi$。

很清楚的，这一定义只能装配在$\sigma[\varphi]$没有被定义[2]的系统中。

另一个更新语义是累积的必要条件是：

幂等律：对每个状态σ和句子φ，$\sigma[\varphi]\Vdash\varphi$

一眼看上去这条规律根本不用说。“用φ来更新你的状态”当然至少意味着“以接受φ的形式来改变你的状态”。然而，仍然有句子不能成功的进行这种更新。像“这个句子是假的”这种悖论句就是这样。如 Groeneveld[1994]中所展示的，这样的句子所具有的悖论性在于，每次当你试图接受它给出的信息时，你都不得不改变自己的想法。

第三个值得看一下的限制条件是保持律。

保持律：如果$\sigma\Vdash\varphi$且$\sigma\leqslant\tau$，则$\tau\Vdash\varphi$。

非保持句的最清楚的例子可以在出现模态限定词“想必(presumably)”、“大概(probably)”、“必须（must）”、“也许（may）”或“或许（might）”[2]的句中找到。考虑以下两个序列。第一个不会引起任何问题，但第二个则会有。

有人在敲门…也许是约翰。…是玛丽。

有人在敲门…也许是约翰。…是玛丽…也许是约翰。

解释：人的期望被事实驳回是非常正常的——这是第一个序列中的情况。但一旦你知道了一些事情，还假装你还期望另一些事情就会有点傻，这是第二个序列所表达的情况。

动态研究方向的一个好处是这些差异能被交代出来。这种设置使我们能够处理（整个

[2] 译者注：might 在本文中翻译成或许，用来表示程度比 may（也许）稍弱。

文本的）句子序列。令φ_1=“有人在敲门”，φ_2=“也许是约翰”，φ_3=“是玛丽”。如果我们想，我们可以对任何状态σ比较$\sigma[\varphi_1]$ $[\varphi_2]$ $[\varphi_3]$和$\sigma[\varphi_1]$ $[\varphi_2]$ $[\varphi_3]$ $[\varphi_2]$来看看是否有些不同。

还有两个更重要的限制：

加强性（strengthening）：$\sigma\leqslant\sigma[\varphi]$

单调性：如果$\sigma\leqslant\tau$，则$\sigma[\varphi]\leqslant\tau[\varphi]$。

关于这些我们还可以说得更多。但现在，我们只给出：

命题 1.2 一个更新系统$<\mathscr{L}, \Sigma, [\ ]>$是累积的，当且仅当

(i) Σ是一个更新格，其中[ ]是全(total)的。且

(ii)满足幂等律、保持律、单调性和加强性。

1.2 有效性概念

有不同的有效性概念。和我最相关的是这个：

- 一个论证是有效的，当且仅当用前提集序列$\psi_1,\ldots,\psi_n$按序更新最小状态 **0**，得到的信息状态中结论φ是被接受的。形式化如下：

$$\psi_1,\ldots,\psi_n \Vdash_1 \varphi \text{当且仅当 } 0[\psi_1]\ldots[\psi_n]\Vdash\varphi。$$

一个更一般的有效性概念如下：

- 一个论证是有效$_2$的，当且仅当由前提集序列$\psi_1,\ldots,\psi_n$按序更新任何状态σ，结论φ在所得的信息状态下成立。形式化如下：

$$\psi_1,\ldots,\psi_n \Vdash_2 \varphi \text{当且仅当对所有的}\sigma，\sigma[\psi_1]\ldots[\psi_n]\Vdash\varphi。$$

下一概念则和经典的最接近：

- 一个论证是有效$_3$的当且仅当人们如果不能接受结论则不能接受全部的前提。更形式化的表达如下：

$$\psi_1,\ldots,\psi_n \Vdash_3 \varphi \text{当且仅当对所有满足的}\sigma\Vdash\psi_1,\ldots,\psi_n\text{的}\sigma，\sigma\Vdash\varphi。$$

命题 1.3 对所有的累积更新系统，如下成立：

$$\psi_1,\ldots,\psi_n \Vdash_1 \varphi \text{当且仅当} \psi_1,\ldots,\psi_n \Vdash_2 \varphi \text{当且仅当} \psi_1,\ldots,\psi_n \Vdash_3 \varphi。$$

一般的，这三个概念并不等同。注意，有效性$_3$是单调的：如果一个有前提$\psi_1,\ldots,\psi_n$和结论φ的论证是有效的，则如果你在$\psi_1,\ldots,\psi_n$上加上更多的前提，仍保持有效$_3$。有效性$_2$至少是左单调的：

$$如果\psi_1,\ldots,\psi_n \Vdash_2 \varphi，则\chi,\psi_1,\ldots,\psi_n \Vdash_2 \varphi。$$

有效性 $_1$ 既不左单调也不右单调。但很容易验证，这一概念符合以下矢列单调性（sequential Monotony）：

$$如果\psi_1,\ldots,\psi_n \Vdash_1 \varphi 且\psi_1,\ldots,\psi_n,\varphi,\theta_1,\ldots,\theta_k \Vdash_1 \chi，则\psi_1,\ldots,\psi_n,\varphi,\theta_1,\ldots,\theta_k \Vdash_1 \chi。$$

进一步的，有效性 $_1$ 符合以下版本的切割消去律，我们应将其称为矢列切割（sequential cut）：

$$如果\psi_1,\ldots,\psi_n \Vdash_1 \varphi 且\psi_1,\ldots,\psi_n,\varphi,\theta_1,\ldots,\theta_k \Vdash_1 \chi 则\psi_1,\ldots,\psi_n,\theta_1,\ldots,\theta_k \Vdash_1 \chi。$$

如果满足幂等律，则有效性是自返的。

$$\psi_1,\ldots,\psi_n,\varphi \Vdash_1 \varphi$$

矢列单调性、矢列切割、自返性完全描述了更新系统中幂等律成立的有效性 $_1$ 概念的结构性质（参考 Benthem[1991]的方法来证明）。

1.3 总结

下一节将讨论一个简单的非累积更新系统。它刻画认知可能性算子“或许 (might) ”的动态。同时更多的术语会被引入。具体的，累积的命题的更新和非经典的测试（test）会被区分开来。

§3 中会研究一个稍微复杂点的系统，包括规则“通常是这样…”和由它引起的预期“想必是这样…”之间的相互作用。规则是经典的，就像普通的描述句，尽管它们所引起的更新并不是命题（逻辑）的。

§4 是本篇论文的核心。这里，§3 发展的系统被添加限制规则而得以扩充，即，形为“如果…，则通常是这样”的句子，我将展示，这些句子的逻辑特性可由一个一致的限制来解释，这一限制决定什么时候一个规则是可接受的。同时，补充一个适用性标准来解释为什么有时候一个规则会被其它规则所否决。

最终，在§5，我们将看到在§4 中发展出的系统足够丰富到来处理我们过去称之为试金石（bentch-mark）的问题。

这里给出一些例子：在§4 和§5 中发展的系统下，这些论证是有效 $_1$ 的：

前提 1：P 通常是 R
前提 2：*x*是P
结论：想必，*x* 是 R

这一论证在我们知道更多有关对象 *x* 的信息且没有证据表明这些信息与结论相关时，保持有效性 $_1$。因此，下面这个例子中推理仍旧能通过。

前提 1：P 通常是 R
前提 2：x 是 P
前提 3：x是Q
结论：想必，x 是 R

然而，如果在前提 1，2，3 上又吸收“Q 通常不是 R”，则这一论证就不再是有效$_1$的。如果人们知道

前提 1：Q 通常不是 R
前提 2：P 通常是 R
前提 3：x 是 P
前提 4：x 是 Q

则是否可预设“x 是 R”是个开放的问题。清楚的，对象 x 一定是其中一个规则的例外，但没有理由认为它是其中一个例外甚于另外一个。而加入更多的常识规则则可能使这种平衡被破坏。如果，例如，我们加入“Q 通常是 P”作为前提，我们得到以下有效$_1$论证：

前提 1：Q 通常是 P
前提 2：Q 通常不是 R
前提 3：P 通常是 R
前提 4：x 是 P
前提 5：x是Q
结论：想必，x 不是 R

由于“Q 通常是 P”的存在，规则“Q 通常不是 R”比“P 通常是 R”优先了。（如果需要一个精确的例子，取“x 是 P”为“x 是成人”，“x 是 Q”为“x 是学生”，“ x 是 R”为“x 有工作”）。

以上的论证都不是有效$_2$或者有效$_3$的。有效$_2$和有效$_3$的定义都包含了一个状态集上的量化（quantification）。因此当检验一个论证的有效性$_2$和有效性$_3$时，人们必须顾及更多信息被知晓的可能性，而不只是前提所陈述的那些。我们由常识规则得出结论，然而，通常这种得出是因为“缺乏任何相反的信息”，当新信息加入时，这些结论可能被收回。因此，在评估一个常识论证时，明确哪些信息是已知的很重要。这是为什么我会锁定有效性$_1$。

动态设置以及随之而来的有效性$_1$概念是本文与其它常识理论不同之处。本理论与其它理论的另一个不同之处在于：由常识得到一个结论的事实可在对象语言中显现。有效性$_1$不是从“P 通常是 R”和“x 是 P”得到“x 是 R”，只是想必是这样。想必的句子是非保持的。因此，这一限制使结论的可废止性更加明确化。在其它理论中，一个结论是由常识规则得出的并没有被标记出来；仅仅在元层次上指出可废止的结论有特殊的地位。

最后，来自优先问题上的想法是本文得以成文的出发点，这一由于常识发生冲突的案例所引起的问题应在该语义层面上有所决断。在有规则“Q 通常是 P”的情况下，规则“Q 通常不是 R”能覆盖规则“P 通常是 R”（见上面最后一个例子）。这是由这些规则的意思所决定的。这不是凌驾在语义之上的规定——像大多数理论中那样——而是要通过解释得到的。

2. 第一个例子：或许（might）

定义 2.1 设 **A** 是有穷多的原子句子组成的集合。以 A 为基础，我们有语言$\mathcal{L}_0^{\mathrm{A}}$和$\mathcal{L}_1^{\mathrm{A}}$，它们都以 **A** 为它们的非逻辑词汇表。$\mathcal{L}_0^{\mathrm{A}}$中有逻辑符号：一元算子¬，二元算子∧和∨，括号) 和 (。$\mathcal{L}_0^{\mathrm{A}}$中只包含以上符号。$\mathcal{L}_1^{\mathrm{A}}$在$\mathcal{L}_0^{\mathrm{A}}$基础上加入一元算子*或许*，符号串$\varphi$是$\mathcal{L}_1^{\mathrm{A}}$中的句子当且仅当存在$\mathcal{L}_0^{\mathrm{A}}$中的句子$\psi$，使得或者$\varphi=\psi$，或者$\varphi=$*或许*$\psi$。

以下，在元语言中，我们用“p”，“q”，“r”等表示原子公式。这些不同的元变元指称不同的原子公式。符号“φ”，“ψ”和“χ”用来表示任意公式。

“或许”背后的想法是：如果φ和一个主体的知识（集）是一致的，那么这个主体就不得不接受*或许* φ——或甚至将其当做知识。否则，*或许*φ就被拒绝掉。

为了将这一想法在数学模型中表示，我们需要表达主体的知识。以下，一个知识状态[3]σ由一个**A**的子集的集合给出。直观上，**A**的一个子集w——或者我们应称其为一个可能世界——是σ的一个元素，如果状态中的所有主体知道，w 或许给出一个事实的正确图画——给定该主体的信息，不排除w中原子公式都真，其它假的可能性。

A 的幂集决定了优先可能性的空间：如果主体恰巧什么都不知道，则 **A** 的任何子集都可能正确的描述现实（世界）。当主体的知识增加时，σ收缩，直到σ包含单独一个 **A** 的子集。这样，该主体的知识是完全的。因此，知识的增长可以被理解为一个消去的过程。

定义 2.2 设 W 是原子公式集的幂集。

(i) σ是一个*信息状态*，当且仅当$\sigma\subseteq W$；

(ii) **0**，*最小状态*(minimal state)，是由 W 给出的信息状态；

1，*缪状态*(absurd state)，是由空集给出的信息状态；

(iii) 对任意两个状态σ和τ，$\sigma+\tau=\sigma\cap\tau$。

注意$\sigma\leqslant\tau$当且仅当$\tau\subseteq\sigma$。

信息状态的概念是依赖于语言的：不同的原子句集合导致不同的可能信息状态集。在定义中就隐含了这一点。因此写为 **A-**信息状态，**A-**最小信息状态更为精确。我将偶尔使用后一术语，尤其是所进行的证明与具体使用哪一个语言没有太大关系时。

定义 2.3 给定 **A**。对$\mathcal{L}_1^{\mathrm{A}}$中的每个句子$\varphi$和陈述$\sigma$，$\sigma[\varphi]$如下：

原子：　$\sigma[\mathrm{p}]=\sigma\cap\{w\in W|\mathrm{p}\in w\}$

¬：　$\sigma[\neg\varphi]=\sigma\sim\sigma[\varphi]$

∧：　$\sigma[\varphi\wedge\psi]=\sigma[\varphi]\cap\sigma[\psi]$

∨：　$\sigma[\varphi\vee\psi]=\sigma[\varphi]\cup\sigma[\psi]$

或许:　σ[或许φ]=σ，如果$\sigma[\varphi]\neq 1$

　　σ[或许φ]=1，如果$\sigma[\varphi]=1$

更新从句对每一个句子φ和每个状态σ告知当状态σ中某主体接受φ时怎样改变。如果$\sigma[\varphi]\neq 1$，φ在σ中*可接受*。如果$\sigma[\varphi]=1$，φ在σ中*不可接受*，如果$\sigma[\varphi]=\sigma$，φ在σ中*被接受*。

这些概念更规范而不仅仅是描述：如果$\sigma[\varphi]=1$，则状态σ中的一个主体*不应该*接受φ。如果$\sigma[\varphi]=\sigma$，状态σ中的主体不得不接受φ。一个拒斥这样做的主体自觉或不自觉的在打破¬，∧，∨，或许等的使用惯例。

另一个需要时刻记在心里的问题是：这些概念与真假概念没多大关系或者没关系。非常有可能$\sigma[p]=1$，而事实 p 为真；或$\sigma[p]=\sigma$，而事实为 p 假。假设 p 事实真，则$\sigma[p]=1$。根据以上给定的术语，p 对状态σ中的一个主体不是可接受的。这是不是意味着状态σ中的个体必须拒绝接受 p，甚至当他/她面对着事实的时候？当然不是。句子 p 在状态σ是不被接受的。因此，主体应*修改*σ使得 p 变成可接受的。在定义 2.3 中，我们不处理修改（revision）：更新从句并没有说对任意句子φ，如果φ在一个状态中是不被接受的，则这个状态必须被修改，使得在结果中φ可以被接受。

它们止步于：很清楚可以看到，如果包含φ的信息吸收到σ中，不一致就会出现的地方。

注意，对每个句子φ，$\mathbf{1}[\varphi]=1$。因此，在缪状态下，每个句子都是被接受的，但没有句子是可接受的。这解释了尽管我们不处理修改，幂等律仍旧成立：即使一个句子φ在σ中不是可接受的——即使你*不应该*接受φ——用φ更新σ的结果是φ在其中*被接受*的一个信息状态。

尽管我们不处理信念修改，一个句子在一个阶段被接受，而之后又被拒斥是经常发生的。改变不是非-保持的唯一可能的来源，测试（testing）是另一个。这里，形式为或许φ的句子是一个例子。如定义中所说，当被告知或许φ时，所有你能做的就是同意或不同意。如果在你的信息状态σ中接受φ，你必须接受或许φ。如果在σ中不被接受φ，则或许φ也不被接受。很清楚的，形为或许φ的句子提供了σ上的一个测试，而不是吸收新信息。这种测试的结果可以先是肯定的，然后是否定的。在最小状态下，你不得不接受“或许在下雨”，但只要你知道没有下雨，“或许在下雨”不得不马上被拒掉。

定义 2.4 一个句子序列$\psi_1,\ldots,\psi_n$是*一致的*，当且仅当，存在一个信息状态，使得$\sigma[\psi_1]\ldots[\psi_n]\neq 1$。

再一次的，因为信息状态集随着$\psi_1,\ldots,\psi_n$组成的语言的非逻辑语汇的变化而变化，更精确的说法应该是 **A**-一致性。下面的引理和命题表明，前缀 **A** 可以被去掉。

引理 2.5 令 $\mathbf{A}\subseteq\mathbf{A}'$，对每个 **A**-状态$\sigma$，我们关联 **A′**-状态$\sigma^{*}=\{w\subseteq\mathbf{A}' \mid w\cap\mathbf{A}\in\sigma\}$，对每个 **A′**-状态，我们关联 **A**-状态$\sigma^{\circ}=\{w\subseteq\mathbf{A} \mid w=v\cap\mathbf{A}$，对某些 $v\in\sigma\}$。对每个$\mathcal{L}_1^{A}$中的φ，以下

成立：

(i)如果σ是一个 **A**-状态，则$\sigma[\varphi]^*=\sigma^*[\varphi]$

(ii) 如果σ, τ是 **A**-状态且$\sigma\neq\tau$，则$\sigma^*\neq\tau^*$

(iii) 如果σ是一个 **A′**-状态，则$\sigma[\varphi]^\circ=\sigma^\circ[\varphi]$

(iv) 如果σ是一个 **A′**-状态，且$\sigma[\varphi]\neq\sigma$，则$\sigma[\varphi]^\circ\neq\sigma^\circ$

命题 2.6 设 $p_1,\dots, p_k$ 是出现在$\psi_1,\dots, \psi_n, \varphi$中的原子句。假设$\{p_1,\dots, p_k\}\subseteq$**A** 且$\{p_1,\dots, p_k\}\subseteq$**A′**。

(i) 论证$\psi_1,\dots,\psi_n/\varphi$是 **A**-有效$_1$的，当且仅当它是 **A′**-有效$_1$的；

(ii) $\psi_1,\dots,\psi_n$ 是 **A**-一致的，当且仅当$\psi_1,\dots,\psi_n$ 是 **A′**-一致的。

假设 $p_1,\dots, p_k$ 是论证$\psi_1,\dots,\psi_n/\varphi$的原子句。由命题 2.6 可得$\psi_1,\dots,\psi_n/\varphi$是否有效是不依赖于语言的，也本该如此。实际上，在考虑这一问题（是否有效）的答案时，我们可以限制在由 **A**$=\{p_1,\dots, p_k\}$生成的状态集上。因为这只有有穷多个，该逻辑是可判定的。

今后，我将省略有效性$_1$和$\Vdash_1$的下标“1”。下面的例子展示了前一节中的某些关键点。

例 2.7

(i) 或许$\neg p, p$ 是一致的；

p, 或许$\neg p$ 不一致；

(ii) 右-单调性不成立：或许$\neg p\Vdash$或许$\neg p$，但或许$\neg p, p\Vdash$或许$\neg p$ 不成立；

(iii) 左-单调性也不成立：$\Vdash$或许 p，但并非$\neg p\Vdash$或许 p。

对或许的逻辑性质的系统化研究要留待其它机会了。以下给出一些初步的观察（结论），它们将在下一节中起到作用。

引理 2.8 设σ和τ是信息状态，φ是$\mathcal{L}_1^A$的句子。

(i) $\sigma\leqslant\sigma[\varphi]$；

(ii) $\sigma[\varphi][\varphi]=\sigma[\varphi]$；

(iii) 如果$\sigma\leqslant\tau$，则$\sigma[\varphi]\leqslant\tau[\varphi]$；

(iv) 如果φ是$\mathcal{L}_0^A$中的句子，则以下成立：

如果$\sigma\leqslant\tau$且$\sigma\Vdash\varphi$，则$\tau\Vdash\varphi$。

加强性，幂等律，单调性和保持性在$<\mathcal{L}_0^A, \Sigma, [\ ]>$下成立。这样，系统$<\mathcal{L}_0^A, \Sigma, [\ ]>$是累积的：我们能给$\mathcal{L}_0^A$中每个句子$\varphi$一个静止的含义，$\mathbf{0}[\varphi]$。用$\varphi$来更新任意状态$\sigma$，归结为取$\sigma$和 $\mathbf{0}[\varphi]$的交点。以下，当我们处理$\mathcal{L}_0^A$中的句子φ时，我们称 $\mathbf{0}[\varphi]$被φ所命题表达，记为$\|\varphi\|$（替代 $\mathbf{0}[\varphi]$）。

作为一个静止设置的起点，可证明有：

$$\|p\|=\{w\in W|p\in w\}$$

$$\|\neg\varphi\|=W\sim\|\varphi\|$$

$$\|\varphi\wedge\psi\|=\|\varphi\|\cap\|\psi\|$$
$$\|\varphi\vee\psi\|=\|\varphi\|\cap\|\psi\|$$

给定这些，则顺理成章的有：对$\mathcal{L}_0{}^{A}$的句子：$\psi_1,\ldots,\psi_n\Vdash\varphi$当且仅当论证$\psi_1,\ldots,\psi_n/\varphi$在经典逻辑中有效。

系统$<\mathcal{L}_1{}^{A}, \Sigma, [\ ]>$不是累积的。形为或许$\varphi$的句子不是保持的；它们不表达一个命题；它们的信息内容并非独立于情境的（context）。如果你获知了$\mathcal{L}_0{}^{A}$中的一个句子φ，你获知现实世界是表达式φ成立的世界之一：现实世界是一个φ-世界。但说"或许φ-世界"是没有意义的。如果φ或许真的，这不是该世界的性质，而是你关于这个世界的知识。

3. 有例外的规则

在上一节中，我们研究了一个简单的更新进程。主体唯一可获得的信息是真正的事实。在本节中，我们关心一个更复杂的进程：主体不仅能获知事实上成立的命题，也能获知通常成立的命题。在此基础上，它们将能决定是否——鉴于信息在手——一个给定的命题想必成立。

定义 3.1 **A** 和$\mathcal{L}_0{}^{A}$同§2。语言$\mathcal{L}_2{}^{A}$有 **A** 为非逻辑词汇表，逻辑词汇表中添加两个一元算子：通常和想必。符号串φ是$\mathcal{L}_2{}^{A}$中的句子当且仅当存在$\mathcal{L}_0{}^{A}$中的句子ψ，使得$\varphi=\psi$或者$\varphi=$通常ψ，或$\varphi=$想必ψ。

以下，形为通常φ的句子称为（常识）规则。为了描述它们对主体心智状态的影响，与上节相比，我们必须给一个信息状态以更多的结构。我们想刻画两点：一个主体的知识以及一个主体的期望。我们想描述当一个主体的知识增加时，他/她的期望是怎样被调整的。一种方法是把一个状态σ当做一个对$<\varepsilon, s>$，其中 s 是可能世界集的子集，与上节中起大体相同的作用；它代表主体的事实知识。集合ε表示主体的规则知识。

定义 3.2 W如前，则ε是 W 上的（期望）模式当且仅当ε是 W 上的自返和传递关系。

关系ε以如下方式编码主体所熟知的规则：设 P 是某一主体认为通常是这样的所有命题组成的集合，我们称$<w, v>$是该主体的期望模式ε的一个元素，如果对 P 中的每个命题，如果该命题在 v 中成立，则它在 w 中也成立，换言之，w 确认所有 P 中的 v 确认的规则，甚至更多。

取代"$<w, v>\in\varepsilon$"，我们经常写成"$w\leqslant_\varepsilon v$"。如果 $v\leqslant_\varepsilon w$ 和 $w\leqslant_\varepsilon v$ 都成立，我们记作"$w\cong_\varepsilon v$"。很清楚的，$\cong_\varepsilon$是一个等价关系。如果有 $v\leqslant_\varepsilon w$，但并非 $w\leqslant_\varepsilon v$，我们记作"$v<_\varepsilon w$"，并称 v 比 w 的例外少。

定义 3.3 设ε是 W 上的模式

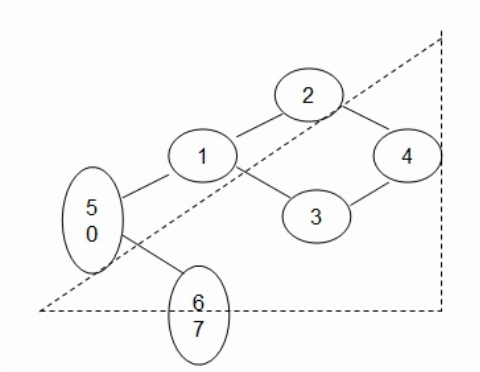
(i) w 是ε中的正常世界，当且仅当 $w\in W$，且对任意 $v\in W$，$w\leqslant_\varepsilon v$；

(ii) $\mathbf{n}\varepsilon$是ε中所有正常世界的集合；

(iii) ε是一致的当且仅当 $\mathrm{n}\varepsilon \neq \emptyset$。

再一次的，设 P 是某一主体认为通常是这样的所有命题的集合。假设$\emptyset \notin P$（规则通常φ可接受的一个必要条件是由φ表示的命题至少在一个世界中为真）。给定这一点，(iii)是说一个模式ε是一致的，当且仅当存在至少一个可能世界，其中所有 P 中命题都成立。要求模式在这一意义上一致是合理的。如果每件事是正常的甚至都不可想象，则有些地方出了问题。这当然不意味着，每件事都在现实中是正常的，或人们必须在所有情境下期待所有事都正常。期望事情比留有余地的数据更正常是不大现实的。

用图例来展示一下总是有帮助的。下图展示了包含三个原子句的语言下的一个状态σ=<ε, s>。



如果两个世界属于同一个$\cong_\varepsilon$-等价类，它们被放入同一个圆/椭圆中。因此，$\cong_\varepsilon$-等价类是{w_1}，{w_2}，{w_3}，{w_4}，{w_0，w_5}和{w_6，w_7}。如果 $w_{\mathrm{i}} <_\varepsilon w_{\mathrm{j}}$，则图中包含一个向右的路径：从$\cong_\varepsilon$-等价类到包含 w_{i} 的等价类到包含 w_{j} 的等价类。例如，若 $w_0 <_\varepsilon w_3$，则没有 $w_2 \leqslant_\varepsilon w_3$ 也没有 $w_3 \leqslant_\varepsilon w_2$。组成 s 的世界由虚线框标出：s={ w_3, w_4, w_6}。正常世界 w_5 和 w_0，不属于 s。因此，一个在状态σ中的主体知道现实世界并非正常的。在可能是现实世界的世界中，w_3 和 w_6 有特殊的地位：它们在下面定义的意义上是极佳的（optimal）。

定义 3.4 设ε是 W 上的模式，且 $s \subseteq W$。

(i) w 在<ε, s>中是极佳的当且仅当 $w \in s$ 且没有 $v \in s$，使得 $v \leqslant_\varepsilon w$；

(ii) $\mathbf{m}_{<\varepsilon, s>}$是所有<$\varepsilon$, s>中的极佳世界的集合。

当有关事情的事实只有部分被知道又必须下决定的情况下，常识规则尤为重要。在这样的情境下，人们必须估计几种可能性：状态<ε, s>中的所有主体知道，s 中的每个元素都可能给出有关事实的一个正确图画。常识的作用是把可能性的范围缩小：一些元素比另一些元素更正常。状态<ε, s>中的主体将假设现实世界要确定尽可能多的正常性标准；想必，它是极佳世界之一。当期望不得不被调整时，并非极佳的那些世界开始变得重要起来。随着人们知识的增加，s 收缩，且在 s 中极佳的世界可能从 s 中消失，而其它的世界可能变成极佳的。

定义 3.5 设ε和ε'是 W 上的模式，$e\subseteq W$

(i) ε'是ε的一个精致化（refinenment）当且仅当$\varepsilon'\subseteq\varepsilon$；

(ii) $\varepsilon\circ e=\{<v, w>\in\varepsilon|$如果 $w\in e$，则 $v\in e\}$；$\varepsilon\circ e$ 是包含命题 e 的ε的一个精致化。

当一个新规则被习得时精致化算子$\circ$就起作用了。想法如下：假设$<v, w>\in\varepsilon$，则在 w 中成立的规则，在 v 中也成立——至少那些在ε中包含的规则如此。

现在，一个新规则加入了：通常φ。有两种可能性：

(i) $\mathbf{n}\varepsilon\cap\|\varphi\|\neq\emptyset$。存在正常的世界，在其中$\|\varphi\|$成立。这样，新规则兼容于$\varepsilon$中所包含的规则：该规则是可接受的。如果它是可接受的，新的模式变成$\varepsilon\circ\|\varphi\|$。即，如果 $w\in\|\varphi\|$ 但 $v\notin\varphi$，则$<v, w>$不得不从ε中移出。给定这个新规则，则不再是 v 确认所有 w 确认的规则了。

(ii) $\mathbf{n}\varepsilon\cap\|\varphi\|=\emptyset$，这种情况下，新规则与$\varepsilon$中原规则不兼容。因此它是不可接受的。

命题 3.6

(i) $(\varepsilon\circ\emptyset)=\varepsilon$

$(\varepsilon\circ W)=\varepsilon$

(ii) $(\varepsilon\circ e)\circ e=\varepsilon\circ e$

(iii) 如果ε是ε'的一个精致化且$\varepsilon'\circ e=\varepsilon'$，则$\varepsilon\circ e=\varepsilon$

(iv) 如果ε是ε'的一个精致化，则$\varepsilon\circ e$ 是$\varepsilon'\circ e$ 的一个精致化。

(ii)、(iii)、(iv)是证明规则满足幂等律，保持律和单调性的基础。

设ε是一个模式，命题 $e\subseteq W$ 是ε中的一个常识当且仅当$\varepsilon\neq\emptyset$且$\varepsilon\circ e=\varepsilon$。下一个命题通过解释上面给出的模式的概念来展示这个术语的是恰当的。

命题 3.7 设ε是 W 上的一个模式。则对任意 $v, w\in W$， $w\leqslant_\varepsilon v$，当且仅当对ε中每个常识 e，$w\in e$ 使得 $v\in e$。

我还始终没有给出一个信息状态的官方的陈述。

定义 3.8 设 W 如前。

(i) σ是一个信息状态，当且仅当$\sigma=<\varepsilon, s>$且满足一下条件之一：

(a) ε是 W 上的一个一致的模式且 s 是 W 的一个非空子集；

(b) $\varepsilon=\{<w, w>|w\in W\}$且 $s=\emptyset$。

(ii) **0**，最小状态，是由$<W\times W, W>$给出的状态；

1，缪状态，是由$<\{<w, w>|w\in W\},\emptyset>$给出的状态。

(iii) 设$\sigma=<\varepsilon, s>$且$\sigma'=<\varepsilon', s'>$是状态。

$\sigma+\sigma'=<\varepsilon\cap\varepsilon', s\cap s'>$，如果$<\varepsilon\cap\varepsilon', s\cap s'>$是一致的；

$\sigma+\sigma'=\mathbf{1}$ ，否则。

注意$<\varepsilon, s>\leqslant<\varepsilon', s'>$当且仅当且 $s'\subseteq s$，且$\varepsilon'\subseteq\varepsilon$。

在最小状态 **0**，没有常识被知道：所有的世界都一样正常。

存在很多对$<\varepsilon, s>$，其中ε是不一致的，或 $s=\emptyset$。这中间只有缪状态 **1** 被官方认定为一个信息状态——想法是，不比状态 **1** 少荒谬的其它不协调的状态可等同于 **1**。

定义 3.9 设$\sigma=<\varepsilon, s>$是一个信息状态，对每个$\mathcal{L}_2{}^A$中的句子φ，$\sigma[\varphi]$被确定如下：

- 如果φ是$\mathcal{L}_0{}^A$中的句子，则
  - 如果 $s\cap\|\varphi\|=\emptyset$，$\sigma[\varphi]=\mathbf{1}$；
  - 否则，$\sigma[\varphi]=<\varepsilon, s\cap\|\emptyset\|>$。
- 如果φ=通常ψ，则
  - 如果 $\mathbf{n}\varepsilon\cap\|\psi\|=\emptyset$，$\sigma[\varphi]=\mathbf{1}$；
  - 否则$\sigma[\varphi]=<\varepsilon\circ\|\psi\|, s>$。
- 如果φ=想必ψ，则
  - 如果 $\mathbf{m}_\sigma\cap\|\psi\|=\mathbf{m}_\sigma$，$\sigma[\varphi]=\sigma$；
  - 否则，$\sigma[\varphi]=\mathbf{1}$。

有关想必φ的规则，类似于或许φ的执行测试的一个邀请：如果由φ表达的命题在σ的所有极佳世界中成立，则句子想必φ一定是可接受的。否则，想必φ是不可接受的——在σ中不可接受。

形为想必φ的句子不意味着传达新信息。通过断言想必φ，说话者做了某种注解：“给定一些常识和事实。我们习得φ将被期望”。接受信息的人也被假定来决定，基于他/她自己的信息，是否接受φ。如果不是这样，将会有这样的讨论：接收信息的人将问，“为什么你认为φ将被期望？”并且在随后的信息改变中，谈话者和接收者可能学到更多新的常识和事实，因此最终两者（说话者和接收者）的期望会相同（要承认，这是有点像田园诗般的画面）。

例 3.10

(i) **0**[通常 p][¬p]≠1

0[通常 p][通常¬p]=1

(ii) 通常 p ⊩ 想必 p

通常 p, ¬p⊮想必 p

通常 p, ¬p ⊩通常 p

(iii) 通常 p, q⊩想必 p

通常 p, q, ¬p⊮想必 p

(iv) 通常 p, 通常 q⊩想必 p

通常 p, 通常 q, ¬p⊮想必 p

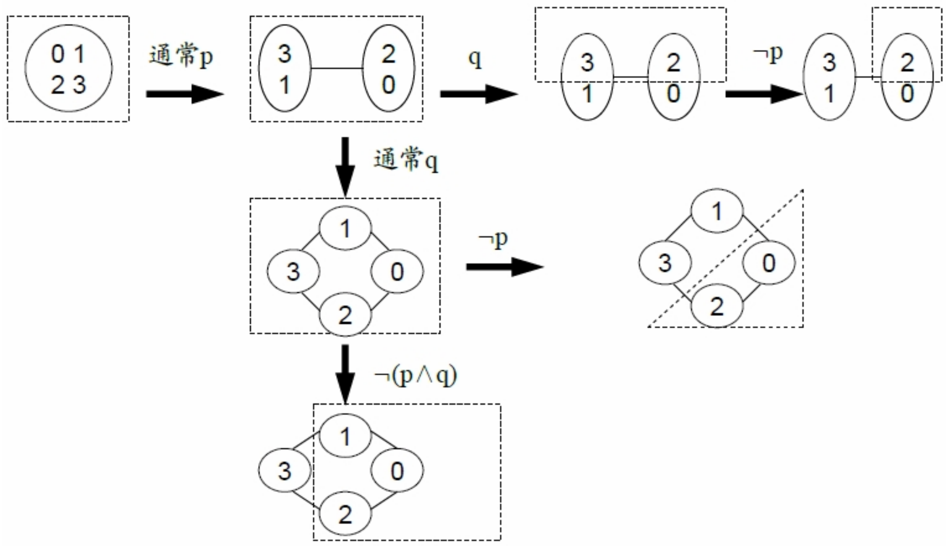
通常 p, 通常 q, ¬p⊩想必 q

(v) 通常 p, 通常 q, ¬(p∧q) ⊮想必 p

通常 p, 通常 q, ¬(p∧q) ⊭想必 q

以上的例子演示了系统的一些重要特征。(i)中的第一个例子表明规则能有例外：一个主体可以先习得通常 p——“通常下雨”——之后发现事实是没下雨。然而，一旦主体接受了“通常下雨”，与之相对的规则“通常不下雨”就不被接受了。

图示描述(ii)、(iii)、(iv)和(v)所描述的例子如下。$W=\{w_0, w_1, w_2, w_3\}$，其中$w_0=\emptyset$，$w_1=\{p\}$，$w_2=\{q\}$，$w_3=\{p, q\}$。(ii)中的前两个例子表明，形为想必φ的句子是不保持的。如果“通常下雨”是一个规则且这是所有你知道的，你可以假定现在正在下雨。但是一旦你知道事实是现在没有下雨，继续保持原规则是愚蠢的。注意，这不意味着你不得不放弃这个有问题的规则。今天的天气可能是个例外，明天想必又正常了。[4]即使人们从一个规则得到的后承不是保持的，该规则自己是[5]。



(iii)和(iv)中的例子的重点在于：已经接受一个规则通常 p，如果你已知的其它信息与 p 无关时——或至少不知道与 p 有关，你也许期待有 p。因此，如果规则是“通常下雨”，而在此之上，你还不知道的就是刮东风，你也许会假定现在正在下雨。（下一节我们将看到，当我们得知刮东风通常意味着天气干燥时会发生什么）。

例(iv)展示了，形为通常φ的句子所说的远远多于φ在所有正常的世界里成立。它导致一个一般的偏好：更倾向于φ成立的世界，而不是不成立的世界。因此，如果现实世界在一方面是个例外，人们仍可假定它在其它方面是正常的。

例(v)表明的是，有时人们会面对困境。如果你倾向于 p 成立的世界而不是¬p 成立的世界，同时倾向于 q 成立的世界而不是¬q 成立的世界。那么，当你不能同时让两者成立时，这种选择是困难的。或者用如下形式来表达：状态 **0**[通常 p][通常 q][¬(p∧q)]是含混的（ambiguous）。

定义 3.11 设$<\varepsilon, s>$是一个信息状态。

(i) **m** 是$<\varepsilon, s>$中一个极佳集，当且仅当，存在$<\varepsilon, s>$中的某个极佳世界 w，使得 $\mathbf{m}=\{v\in s|v\cong_{\varepsilon}w\}$；

(ii) $<\varepsilon, s>$是含混的，如果$<\varepsilon, s>$中存在不止一个极佳集。

我将不在这里系统化的研究正常和想必了。

然而，以下的内容是基础的。

引理 3.12 设φ是$\mathcal{L}_2^{\mathrm{A}}$的句子，$\sigma$和$\tau$是任何状态。

(i) $\sigma\leqslant\sigma[\varphi]$；

(ii) $\sigma[\varphi][\varphi]=\sigma[\varphi]$；

(iii) 如果$\varphi\neq$想必ψ且$\sigma\leqslant\tau$，则$\sigma[\varphi]\leqslant\tau[\varphi]$；

(iv) 如果$\varphi\neq$想必ψ且$\sigma\leqslant\tau$且$\sigma\Vdash\varphi$，则$\tau\Vdash\varphi$。

我们已经看到形为想必φ的句子不是保持的。由于对想必φ的测试通常是先有负面的结果，再有正面的结果，它们还是非单调的。举例说明要注意的，有 $\mathbf{0}\leqslant\mathbf{0}$[p]，但没有 $\mathbf{0}$[想必 p] $\leqslant\mathbf{0}$[p] [想必 p]。

注意，引理 3.12 中的(iii)和(iv)对规则适用。我们可以给通常φ指派一个静态的含义，即，$\mathbf{0}$[通常φ]。考虑用通常φ更新状态σ的进程如下：对σ增加包含在 $\mathbf{0}$[通常φ]中的信息。不仅纯粹的描述句携带情境独立的信息，规则也是。

对算子通常的逻辑性质有所洞察的一种方式是将其与真势必然性算子做比较。下面的规律给出后者在模态逻辑的通常的系统中逻辑性质的描述[6]。

必然$\varphi\Vdash\varphi$

必然φ，必然$\psi\Vdash$必然$(\varphi\wedge\psi)$

必然$\varphi\Vdash$必然$(\varphi\vee\psi)$

如果$\Vdash\varphi$，则$\Vdash$必然φ

如果我们用通常替换必然，这些规律中只有第 2 和第 4 条仍保持有效。我们发现：

通常φ，通常$\psi\Vdash$通常$(\varphi\wedge\psi)$

如果$\Vdash\varphi$，则$\Vdash$通常φ

我们已经知道，对通常来说，第一条规律是不成立的。我们有更弱的规律：

通常$\varphi\Vdash$想必φ。

第三条规律也不成立。一般没有

$$\text{通常}\varphi \Vdash \text{通常}(\varphi \vee \psi)\text{。}$$

也许给个例子会更清楚。比较：

——通常下雨。现在没下雨。所以，想必现在在下雪。
——通常下雨或下雪。现在没下雨，所以，想必现在在下雪。

直觉上，第一行是不正确的。更形式化的说法是，它是非有效的。

$$\text{通常 } p, \neg p \nVdash \text{想必 } q$$

第二行看起来是正确的，形式化一下我们发现

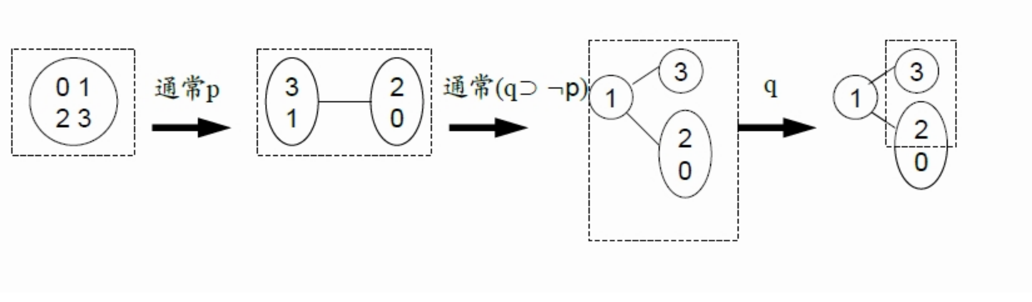
$$\text{通常}(p \vee q), \neg p \Vdash \text{想必 } q$$

上面的例子还表明为什么一个主体可能接受通常p，同时却拒绝接受通常(p∨q)。后者暗示了当p为假时，我们可预期什么，前者则没有。一个主体可以同意通常p成立，但不同意当p为假时，q比¬p更被期望。[7]

4. 关于例外的规则

上面的系统在表达力上是有欠缺的。对于有偶然的例外的一般性规则，它是可以处理的——“通常下雨，但今天不下”——但对于非偶然的例外就不行了：我们不能说什么时候例外的情境被期望，以及当它们获得时，人们期望什么——“通常下雨，但如果刮东风，气候通常是干燥的。”

这里有一个例子。假设一个主体在状态 **0** 接受规则通常 p——通常下雨。这导致一个对‖p‖成立的世界的整体偏好。现在该主体想制造一个意外：如果‖q‖成立，‖p‖通常不成立——如果有东风，则通常不下雨。问题在于这个例外不能由公式通常(q⊃¬p)来制造。结果应该是在 q-世界的域之中，规则通常 p 是被拒斥的，但事情并非如此。公式通常(q⊃¬p)导致了另一个整体偏好，这一次是对命题‖ q⊃¬p ‖成立的世界。因此，当获知现实世界中‖ q ‖成立，一个含混的情境发生了：存在两个极佳结合，一个是对确证通常 p 的世界，另一个是对确证通常(q⊃¬p)的世界。在以下的图中，w_3={p, q }，w_2={q }，w_1={p }且 w_0=∅。



人们不能把“如果q，则通常¬p”与通常q⊃¬p等同。二元算子“如果…, 则通常…”不能以一元算子“通常…”的形式来定义。[8]

定义 4.1 设 **A** 和$\mathcal{L}_0^{\mathrm{A}}$如§2。语言$\mathcal{L}_3^{\mathrm{A}}$以 **A** 为其非逻辑词汇表，逻辑词汇表中加入一个二元算子→和一个一元算子想必，符号串φ是$\mathcal{L}_3^{\mathrm{A}}$中的句子，当且仅当有$\mathcal{L}_0^{\mathrm{A}}$中公式$\psi$和$\chi$，使得$\varphi=\psi$，或$\varphi=\psi\to\chi$或$\varphi$=想必$\psi$。

$\varphi\to\psi$读作“如果φ，则通常ψ”。形成 “$\varphi\to\psi$”的句子将表示在φ给出的世界的域中，$\|\psi\|$是一个常识。如果这个域是一个可能世界集的子集（proper subset），“$\varphi\to\psi$”被称作限制（restricted）规则。形为通常ψ的一般规则在这里以的$(\psi\vee\neg\psi)\to\psi$缩写被重新引入。

定义 4.2

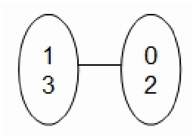
(i) 设 W 如前。W 上的一个框架是一个函数π：π给 W 的每个子集 d 指派一个 d 上的模式πd。

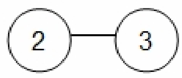
(ii) 设π是 W 的一个框架。$d, e\in W$，命题 e 是πd 上的一个常识，当且仅当 $d\cap e\neq\emptyset$，且$\pi d\circ e=\pi d$。

当我们明确所讨论的框架时，我们说“e 是一个 d-常识”而不是“e 是一个πd 上的常识”。

对于前面给出的例子，结果框架π看起来如下：

$\|$p$\|$是πW 上的一个常识：





$\|\neg$p$\|$是$\pi\|$q$\|$中的一个常识

且对 $d\neq W$ 或 $d\neq\|$q$\|$，$\pi d= d\times d$。

给定定义 4.2，每个 W 上的子集 d 有其自己的模式πd。因此，目前我们的主体可以制造随他们的希望的那么多的例外。当然，不是任意的都可以。如果制造了太多的例外，他们期待的框架会不一致。

4.1 一致

定义 4.3 设π是 W 上的框架，且 $d\subseteq W$。

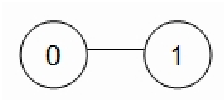
(i) w 是πd 中的一个正常世界当且仅当 $w\in d$ 且对每个 $d'\subseteq d$ 且 $w\in d'$的 d'，如下成立：对每个 $v\in d'$， $w\leqslant_{\pi d'} v$ ；

(ii) $\mathbf{n}\pi d$ 是πd 中所有正常世界的集合；

(iii) π是一致的当且仅当对每个非空 $d\subseteq W$，$\mathbf{n}\pi d\neq\emptyset$。

考虑上面描述的框架。给定定义 4.3，$\mathbf{n}\pi d=\{w_1\}$。因此尽管 w_3 符合一般规则通常 p 是事实，w_3 不被算作πW 中的一个正常世界。考虑如下：通过接受作为‖q‖-常识的‖¬p‖，主体制造了一个例外，在论域‖q‖中的世界从一般规则中被豁免了。因此，像之前那样说 w_3 适合一般规则，像是一种误导：好像在暗示 w_3 首先与这条规则相关，但它不是。w_3 仅仅和更具体的规则 q→¬p 相关，这刚好是一个例外，而世界 w_3 是一个例外从句的例外，我们不打算把“例外的例外”看做正常的。

这里有一个有关框架不一致的简单例子。我们正在面对，一个主体他相信通常下雨，且对有东风的情况制造了一个例外：如果有东风，则通常下雨。在此基础上，主体想制造一个没有东风时的例外：如果没有东风，则通常也不下雨。这就太多了：该主体制造了太多的例外。形式化为：结果框架π'和如上描述的框架π相同，除了现在‖¬p‖是π‖¬q‖中的一个规则。因此$\pi'\{w_0, w_1\}$看起来如下：



但这意味着 $\mathbf{n}\pi W=\emptyset$。框架$\pi'$是不一致的。

定义 4.4 设 W 如前。

(i) σ是一个信息状态当且仅当$\sigma=<\pi, s>$且满足以下条件之一:

(a) π是 W 上的一个一致框架，s 是 W 上的非空子集；

(b) π是框架$<\iota, \emptyset>$，其中对每个 $d\subseteq W$，$\iota\, d=\{<w, w>|w\in d\}$。

(ii) $\mathbf{0}=<\vartheta, W>$，其中对每个 $d\subseteq W$，$\vartheta d=d\times d$。

$\mathbf{1}=<\iota, \varphi>$。

(iii) 设$\sigma=<\pi, s>$和$\sigma'=<\pi', s'>$是状态，π''是框架且对每个 d，$\pi''d=\pi d\cap\pi' d$。这样

$\sigma+\sigma'=<\pi'', s\cap s'>$，如果$<\pi'', s\cap s'>$是一致的；

$\sigma+\sigma'=\mathbf{1}$ ，否则。

这些定义和前一节中相应的定义的不同之处（定义 3.8）都归因于我们现在不是只处理一个模式，而是处理一个模式的框架。

用一个新规则来更新一个信息状态是一个精致化的事件，和从前一样。如果状态$\sigma=<\pi, s>$中的一个主体决定接受$\varphi\to\psi$，则模式π‖φ‖将不得不被‖ψ‖所精致化。不过当然，如果用‖ψ‖精致化π‖φ‖的结果是不一致的，没有主体应该接受$\varphi\to\psi$。

定义 4.5

(i) 设π和π'都是基于 W 的框架。框架π是π'的一个精致化当且仅当对每个 $d\subseteq W$，$\pi d\subseteq\pi' d$。

(ii) 设π是一个框架且 $d, e\subseteq W$。$\pi_{d\circ e}$是π的由以下给出的精致化。

(a) 如果 $d'\neq d$，则$\pi_{d\circ e}d'=\pi\, d'$；

(b) $\pi_{d\circ e}d=\pi d\circ e$。

框架$\pi_{d\circ e}$是在π中由 e 来精致化πd 的结果。

定义 4.6 设$\sigma=<\pi, s>$是一个信息状态。

- $\sigma[\varphi\rightarrow\psi]=\mathbf{1}$, 如果$\|\varphi\|\cap\|\psi\|=\emptyset$或$\pi_{\|\varphi\|\circ\|\psi\|}$是不一致的。
- 否则，$\sigma[\varphi\rightarrow\psi]=<\pi_{\|\varphi\|\circ\|\psi\|}, s>$。

案例$\|\varphi\|\cap\|\psi\|=\emptyset$是特殊的：根据定义 4.2(ii)，$\|\psi\|$在这里不能是$\pi\|\varphi\|$中的常识。依据命题 3.6(i)，$\pi_{\|\varphi\|\circ\|\psi\|}=\pi\|\varphi\|$。这样，$\pi_{\|\varphi\|\circ\|\psi\|}$是一致的——一个技术上的困难。

命题 4.7 设π是一致的，$d, e\subseteq W$，假设 $d\cap e\neq\emptyset$。

$\pi_{d\circ e}$是一致的，当且仅当，不存在 $d'\supseteq d$，使得 $\mathbf{n}\pi d'\subseteq d\sim e$。

把定义和命题组合到一起，我们得到

设$\sigma=<\pi, s>$是一个信息状态。$\sigma[\varphi\rightarrow\psi]$确定如下：

- 如果对某些 $d\supseteq\|\varphi\|$，$\mathbf{n}\pi d\subseteq\|\varphi\|\sim\|\psi\|$，则$\sigma[\varphi\rightarrow\psi]=\mathbf{1}$。
- 否则，$\sigma[\varphi\rightarrow\psi]=<\pi_{\|\varphi\|\circ\|\psi\|}, s>$。

例 4.8

(i) $\mathbf{0}$[通常 p][q→¬p]≠**1**；

(ii) $\mathbf{0}$[通常 p][q→¬p] [¬q→¬p]=**1**；

(iii) $\mathbf{0}$[通常 p][q→¬p] [通常 q]=**1**；

(iv) $\mathbf{0}$[p→q][q→p] [p→r] [q→¬ r]=**1**。

(i)和(ii)前面讨论过。(iii)和(iv)留做练习。

4.2 适用性

设$\sigma=<\pi, s>$是一个状态。框架π包含一个主体在状态σ习得的规则，s 是他/她的事实知识。现在，状态σ中的主体将期望什么？在之前的小节里，我们处理的状态中只包含一个模式ε，这一问题是容易回答的：我们所有要做的是在模式ε下整理出 s 中的极佳世界。这一节中事情就复杂多了。我们要处理一定数量的模式，这些模式不需要在 s 上有相同的影响。

这里的一个重要的概念是适用性：如果你想知道在状态$<\pi, s>$中一个主体所期望的，你不得不整理出π中包含的规则中的哪些适用于 s 内。

定义 4.9 设$\sigma=<\pi, s>$是一致的信息状态且假设 $e_1, \ldots, e_n$ 分别是$\pi d_1,\ldots\pi d_n$ 中的常识。

(i) 一个世界 w 遵守$\{e_1, \ldots, e_n\}$当且仅当对每个 $w\in d_i$ 的 i，$w\in e_i$。 ($1\leqslant i\leqslant n$)。

(ii) 常识集$\{e_1, \ldots, e_n\}$适用于 s 内当且仅当对每个 $d \supseteq s$，存在某个 $w \in \mathbf{n}\pi d$ 使得 w 遵守$\{e_1, \ldots, e_n\}$。

（相比“集合$\{e_1, \ldots, e_n\}$适用于 s 内”的说法，我们通常说“$e_1, \ldots, e_n$ 共同适用于 s 内”）。

为了理解这个过程，我们首先看只处理一个常识的例子。（在这个例子中我们说 d-常识 e，而不是$\{e\}$适用于 s 内）。定义缩减为：

设$<\pi, s>$是一个一致的信息状态，e 是πd 中的一个常识。常识 e 适用于 s 内当且仅当不存在 $d' \supseteq s$，使得 $\mathbf{n}\pi d' \subseteq d{\sim}e$。如果 s 是的一个 d 的子集，则得到一个更特殊的例子。那样，我们说 d-常识 e 适用于 s（而不是适用于 s 内）。

命题 4.10 设π是一个一致的框架。设 e 是一个πd 中的常识，假设 $s \subseteq d$，常识 e 适用于 s 当且仅当存在π的一个一致的精致化π'，使得对每个 $s \subseteq d' \subseteq d$ 的域 d'，e 是一个$\pi' d'$中的常识。

换言之，d-常识 e 适用于 d 的子域 s 仅当 e 是一个 s 和 d 之间的每个域中都可接受的常识。如果 s 和 d 之间有某个域 d'不能一致的精致化 e，则 e 不适用于 s。

例 4.11 对于每个状态$\sigma_i=<\pi_i, s_i>$，我们想知道哪些常识适用于 s_i。

(i) $\sigma_1=\mathbf{0}$[通常 p][q→¬p] [q]；

(ii) $\sigma_2=\mathbf{0}$[通常 p][q→¬p] [q∧r]；

(iii) $\sigma_3=\mathbf{0}$[通常 p][q→¬p] [(q∧r) →p] [q∧r]；

(iv) $\sigma_4=\mathbf{0}$[p→ r][q →(p∧¬r)] [p∧q]。

这里及后面，如果把 p 看成“下雨”， q 看成“有东风”，r 看成“温度低于 15℃”。想象这些案例中的每一个都在讨论一个不同的国家。你对这个国家的气候的所有知识仅仅是通过提到的规则。你对今天天气的了解也只通过提到的描述句。问题在于：你还期望些其它的什么？

例(i)。我们已经知道框架π_1：‖p‖是πW 中的一个常识。且‖¬p‖是π‖q‖中的一个常识。主体的事实知识由 s_1=‖q‖给出。很清楚的，π‖q‖不能一致的被‖p‖精致化。因此，由命题 4.10，‖p‖不适用于 s_1。它们被更具体的‖q‖-常识‖¬p‖所拒绝，‖¬p‖适用于 s_1。

例(ii)。对这个例子，8 种可能性必须被考虑到。除了这些，框架π_2 非常像π_1；其唯一有意思的性质是‖p‖是 W 中的常识，‖¬p‖是‖q‖中的常识。主体的事实知识由 s_2=‖q∧r ‖给出。当π_2‖q‖被‖p‖精致化后，结果是不一致的。既然 $s_2 \subseteq$‖q‖$\subseteq W$，由命题 4.10，有 W-常识‖p‖不适用于 s_2。更特殊的，‖q‖-常识‖¬p‖适用于 s_2。

例(iii)。认识到我们在处理一个三位关系“d-常识 e 适用于 s”是重要的。经常的，第一个主目（argument）会被抑制，但有些时候，我们不能这样做。当我们比较第二和第三个

例子时，这一点变得明显了。上面我们看到σ_2在中 W-常识‖p‖不适用于‖q∧r‖。这没什么不对，然而，如果一个主体除了通常 p 和 q→¬p 还接受规则(q∧r)→p ——作为一个例外从句的例外从句。但即使这么做了，W-常识‖p‖仍不适用于‖q∧r ‖。是更具体的‖q∧r ‖-常识‖p‖适用于其上。

例(i)-(iii)展示了适用性标准怎样施力使得更具体的规则比更一般的规则更优先。然而，像下一个例子所展示的，这不是由它所施力的唯一的事情。

例(iv)。规则 p→ r 和 q→(p∧¬r)并没有谁比谁更具体。然而，在由 p∧q 给出的情境下，只有规则 q→(p∧¬r)被考虑，这是为什么状态σ_4中的一个主体被允许得到以下的结论：

p→ r	如果下雨，通常温度低于 15℃。
q→(p∧¬r)	如果有东风，则通常下雨， 但温度是 15℃或更高。
p∧q	下雨且有东风 。
想必 r	想必，温度是 15℃或更高。

‖p‖-常识‖r‖不适用于 s_4=‖p∧q‖，因为‖q‖⊇ s_4，而 $\mathbf{n}\pi_4$‖q‖⊆‖p‖~‖r‖。‖q‖-常识‖p∧¬r‖适用于‖p∧q‖，因为没有 d⊇‖p∧q‖使得 $\mathbf{n}\pi$d′⊆‖q‖~‖p∧¬r‖。定义 4.9 更关乎常识的集合，而不是单独的常识。从下个例子我们可看出为什么是这样。

例 4.11（续）对每个状态σ_i=<π_i, s_i >，我们想知道哪些常识共同适用于 s_i 内。

(v)　σ_5=**0**[p→r][q→¬r] [p∧q]；

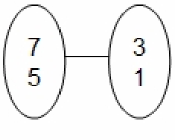
(vi)　σ_6=**0**[q→p] [p → r] [q]；

例(v)。如果下雨，则温度通常低于 15℃，如果有东风，则温度通常为 15℃或更高。下雨且恰巧有东风。会是什么样的温度？以下分析揭示了为什么这里没有更多可说的。

我们处理有 8 个可能世界的集合 W={w_0,…,w_7}，描述如下。集合 s_5={ w_3, w_7}。π_5是如下框架：

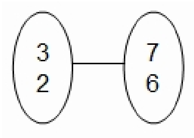
如果 d≠{ w_1, w_3 , w_5, w_7}且 d≠{ w_2, w_3 , w_6, w_7}，$\pi_5 d=d\times d$

索引	世界
0	-
1	p
2	q
3	q, p
4	r
5	r, p
6	r, q
7	r, q, p



$\pi_5\|p\|$看来如右：

$\pi_5\|q\|$是：



因此，如果{ w_1, w_3 , w_5, w_7}$\subseteq d$，$\mathbf{n}\pi d=d\sim${ w_1, w_3}；且如果{ w_2, w_3 , w_6, w_7}$\subseteq d$，$\mathbf{n}\pi d=d\sim${ w_6, w_7}。否则，$\mathbf{n}\pi d=d$。

命题$\|r\|$={ w_4, w_5 , w_6, w_7}作为一个常识在每个s_5={ w_3 , w_7}和$\|p\|$={ w_1, w_3 , w_5, w_7}之间的每个域中是可接受的。因此，$\|p\|$-常识$\|r\|$适用于s_5。同样，我们发现$\|q\|$-常识$\|\neg r\|$适用于s_5。然而，没有的一致的π_5精致化π'使得$\|r\|$={ w_4, w_5 , w_6, w_7}和$\|\neg r\|$={ w_0, w_1 , w_2, w_3}都是π'{ w_3 , w_7}中的常识。这等于说，$\|p\|$-常识$\|r\|$和$\|q\|$-常识$\|\neg r\|$不共同适用于s_5。

命题 4.12 设σ=<π, s>是一个一致的信息状态且假设e_1, …,e_n分别是πd_1,…πd_n中的常识。假设对每个 i ($1\leqslant i\leqslant n$)，$s=d_i$。常识e_1, …,e_n共同适用于s，当且仅当，存在一个一致的π的精致化π'，使得对每个i，对每个满足$s\subseteq d'\subseteq d$域 d′，e_i是π'd′中的一个常识。

这里需注意的最重要的事情是量词的顺序：这是在说“存在一个一致的精致化使得对每个i，e_i是…”而不是“对每个i，存在一个一致的精致化使得e_i是…”。在后例中，e_1, …,e_n中的每个常识分别适用于s，但e_1, …,e_n可能不共同适用。

现在我们转到并不是主体熟悉的所有规则都在s的扩张域中表达常识的例子。

例(vi)。我们将发现 $q\rightarrow p$, $p\rightarrow r$, $q\Vdash$ 想必 r。

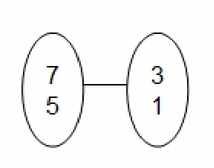
这样的主要原因在于状态 **0** [$q\rightarrow p$] [$p\rightarrow r$] [q]中，$\|q\|$-常识$\|p\|$和$\|p\|$-常识$\|r\|$共同适用于$\|q\|$内。

考虑 W={ w_0,…,w_7}（同例(v)）。

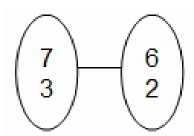
集合s_6={ w_2, w_3 , w_6, w_7}；π_6是如下框架：

如果 $d\neq${ w_1, w_3 , w_5, w_7}且 $d\neq${ w_2, w_3 , w_6, w_7}，$\pi_6 d=d\times d$，

$\pi_6\|p\|$是：



$\pi_6\|q\|$是：



因此，如果$\{w_1, w_3, w_5, w_7\}\subseteq d$，$\mathbf{n}\pi d=d\sim\{w_1, w_3\}$且如果$\{w_2, w_3, w_6, w_7\}\subseteq d$，$\mathbf{n}\pi d=d\sim\{w_2, w_6\}$。否则，$\mathbf{n}\pi d=d$。

如果对每个 d⊇‖q‖，存在某个 $w\in\mathbf{n}\pi d$，遵守‖q‖-常识‖p‖和‖p‖-常识‖r‖，则‖q‖-常识‖p‖和‖p‖-常识‖r‖共同适用于‖q‖内。因为对每个 $d\supseteq$‖q‖，$w_7\in\mathbf{n}\pi d$ 成立。而因为 w_7 是‖q‖中唯一遵守这两个常识的世界，状态σ_6 中的主体将期望现实世界更像 w_7 而不是 w_2、w_3 或 w_6。这意味着主体将期望 p 和 r 都为真。

现在，定义 4.9 背后的基本想法该清楚了。首先，如果一个常识集适用于一个给定的情境 s 内，影响是不遵守这些常识的世界不再被当做正常的 s-世界。第二，从之前的节我们知道，在一个一致的框架中，以下成立：如果一个世界在 s 中不正常，其在 s 的任意扩张域中也不正常。因此，什么时候一个常识集适用于 s 内？如果没有 s 的扩张域 d，（使得）正常 d-世界的 $\mathbf{n}\pi d$ 包含全部不遵守所谈论的常识的全部世界。因为，否则，如果这些常识适用，框架该不一致了。

以上我暗示过几次以下的定义。

定义 4.13 设$\sigma=<\pi, s>$是一个一致的信息状态，且假设 $e_1, \ldots, e_n$ 是$\pi d_1, \ldots \pi d_n$ 中的常识。

(i) 则$\{e_1, \ldots, e_n\}$是一个σ中的最大可适用集，且对每个满足 e_{n+1} 是$\pi_{d_{n+1}}$中的常识，如果 $e_1, \ldots, e_n, e_{n+1}$ 共同适用于 s 内的每个 e_{n+1} 和 d_{n+1}，存在某个 $i\leqslant n$，$e_{n+1}= e_i$ 和 $d_{n+1}= d_i$，成立。

(ii) 一个世界 w 在σ中是极佳的当且仅当 $w\in s$ 且 w 遵守一个最大可适用常识集。极佳世界集由 $\mathbf{m}_\sigma$表示。

(iii) σ [想必ψ]由以下确定：

- 如果 $\mathbf{m}_\sigma\cap\|\psi\|= \mathbf{m}_\sigma$，则$\sigma$ [想必ψ]=σ。
- 否则，σ [想必ψ]=1。

非常有可能出现不止一个最大可适用常识集。如果是这样，该状态被称为含混的。

命题 4.14 设$\sigma=<\pi, s>$是一个一致的信息状态。设每个πd 如下给出：$\pi d=(d\times d)\circ(e_d)_1\ldots\circ(e_d)_m$。

则 w 是σ中极佳的当且仅当 $w\in s$ 且 w 遵守具有以下性质的常识集 D：

(i) D 中每个元素等（价）于某个$(e_d)_i$；

(ii) D 适用于 s 内；

(iii) 对每个使 D$\cup\{(e_d)_i\}$适用于 s 内的$(e_d)_i$，$(e_d)_i\in$D 成立。

假设你不得不弄清楚一个形为$\varphi_1\rightarrow\psi_1, \ldots, \varphi_n\rightarrow\psi_n, \chi_1, .., \chi_m$/想必$\theta$的论证是否是有效的。那么你不得不做的是确定状态$\sigma=\mathbf{0}[\varphi_1\rightarrow\psi_1]\ldots[\varphi_n\rightarrow\psi_n][\chi_1]..[\chi_m]$中的极佳世界的集合。定义 4.13 在说为了这样做，你不得不确定所有σ中的最大适用常识集。

命题 4.14 促成了这个工作：你不必考虑在域$\|\varphi_1\|, \ldots, \|\varphi_n\|$中明确给出的常识

$\|\psi_1\|,\ldots,\|\psi_n\|$之外的常识。所有你要做的就是确定适用于$\|\chi_1\|\cap\ldots\cap\|\chi_n\|$之内的$\{\|\psi_1\|,\ldots,\|\psi_n\|\}$的最大子集。

给定命题 4.14，很容易确定例 4.11 中状态$\sigma_1,\ldots,\sigma_6$中的极佳世界集。这样，我们发现：

(i) 通常 $p, q\to\neg p, q \Vdash$ 想必$\neg p$；

(ii) 通常 $p, q\to\neg p, q\wedge r\Vdash$ 想必$\neg p$；

(iii) 通常 $p, q\to\neg p, (q\wedge r)\to p, q\wedge r \Vdash$ 想必 p；

(iv) $p\to r, q\to(p\wedge\neg r), p\wedge q \Vdash$ 想必$\neg r$；

(v) $p\to r, q\to\neg r, p\wedge q$ 想必$\nVdash r$；

$p\to r, q\to\neg r, p\wedge q\nVdash$ 想必$\neg r$；

(vi) $q\to p, p\to r, q\Vdash$ 想必 r。

5. 比较

到目前为止，我们已经考虑了把语言$\mathcal{L}_3^A$当做命题语言，但我们可以给出它的一个谓词逻辑解释。把 p, q 等看做单子谓词，而不是原子句。每个这样的谓词指定一个性质，且每个$\mathcal{L}_0^A$的良表达（well formed）指定一个性质的布尔组合。把 *W* 看做可能对象（object）的集合，而不是可能世界集。一个可能对象 $i\in W$ 有原子 p 表达的性质，当且仅当 $p\in i$。注意不同的可能对象有不同的性质。因此，把 *W* 的元素称为对象的可能类型 (type) 也许更精确：在现实中可能有不止一个或者没有对象满足 *W* 中一个给定的可能对象的描述。

如前，状态$<\pi, s>$中的集合 *s* 代表主体的知识，只是现在它不是有关现实世界的主体知识，而是有关某个现实对象的。由$\mathcal{L}_0^A$ 中公式φ可知这个对象有φ表达的性质（并没有在φ中被明确提到)。

现在，模式πd 中的一个常识是一个性质——有性质 *d* 的对象通常具有的性质。既然φ-世界（由φ表达的命题成立的世界）已经变成φ-对象（具有φ表达的性质的对象）了，“$\varphi\to\psi$”可以被读作“φ-对象通常是ψ-对象”，而不是“φ-世界通常是ψ-世界”。

这里重复之前我曾说过的一件事：在现实中可能有一个或者没有对象满足一个给定可能对象的描述。期望框架是概念上的框架。因此，如果一致性条件要求 $\mathbf{n}\pi d\neq\emptyset$，这恰恰意味着对 *d* 中的对象来说：有所有该对象通常有的性质是可能的。这不意味着这样的对象必须真的存在。非常有可能在现实中没有对象满足 $\mathbf{n}\pi d$ 中任何个体都有的描述。可能是每个和每一个现实的鸟都缺乏一个或更多鸟通常有的性质，或者根据规则，或者是偶然的。每个鸟在某一方面都是不正常的可以是个事实。但这不能是个规则。如果你希望一个系统中句子“鸟通常是不正常的”是可接受的，那你最好去别处看看。

看看上一节中通过谓词逻辑类来处理的例子，你将看到一些老“相识“。例如，例 4.11（v），也可以作为著名的**尼克松两难困境**（Nixon Dilemma）的形式化：

贵格会成员通常是和平主义者；	$p\to r$
共和党通常不是和平主义者；	$q\to\neg r$

尼克松即是共和党又是贵和会成员。 $p\wedge q$

如我们所见，关于尼克松是不是和平主义者，从这些前提得不出结论，即使是暂时的结论也得不出。

同样著名的还有以下的例子，这个例子我们之前没有讨论过。

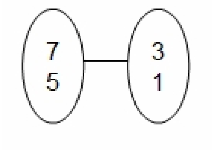
成年人通常是有工作的；	$p\rightsquigarrow r$
学生通常没有工作；	$q\rightsquigarrow\neg r$
学生通常是成年人；	$q\rightsquigarrow p$
约翰是一个学生；	q
想必，约翰是一个成人且没有工作。	想必$(p\wedge\neg r)$

这一论证是有效的。要究其原因，我们不得不确认状态

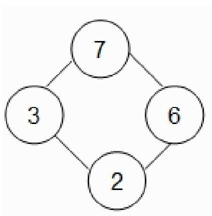
$$\mathbf{0}[p\rightsquigarrow r][q\rightsquigarrow\neg r][q\rightsquigarrow p][q]=\sigma=<\pi, s>。$$

设 W 定义同例 4.11(v)。则 $s=\{w_2, w_3, w_6, w_7\}$。对π，我们发现：如果 $d\neq\{w_1, w_3, w_5, w_7\}$且 $d\neq\{w_2, w_3, w_6, w_7\}$，$\pi d=d\times d$

$\pi\|p\|$可以被描述成：



$\pi\|q\|$为：



由于 $\mathbf{n}\pi s=\{w_3\}\subseteq\|p\|\sim\|r\|$，$\|p\|$-常识$\|r\|$不适用于 s 内。其它规则适用，这意味着$\sigma\Vdash$想必$(p\wedge\neg r)$。

定义 5.1 设$\sigma=<\pi, s>$是一个状态。

(i) σ中包含的事实信息如下给出

$$\{\varphi|\varphi\text{是}\mathcal{L}_0^A\text{中的句子使得 }s\subseteq\|\varphi\|\}。$$

(ii) $\mathcal{L}_0^A$中句子集Δ是σ中包含的事实信息的*扩充*，当且仅当存在某个最大适用常识集E，使得$\Delta=\{\varphi|\{w\in s\ |w\text{ 遵守 }E\}\subseteq\|\varphi\|\}$。

一个比较这里发展出的理论和其它理论的方法是：比较它们不得不提到的扩充。例如我们有：

$$\varphi_1\rightsquigarrow\psi_1, \ldots, \varphi_n\rightsquigarrow\psi_n, \chi_1,..,\chi_m\Vdash\text{想必}\theta\text{当且仅当}\theta\text{属于}$$

$$\mathbf{0}[\varphi_1\rightsquigarrow\psi_1]\ldots[\varphi_n\rightsquigarrow\psi_n][\chi_1]..[\chi_m]\text{中的事实信息的每一个扩充。}$$

换言之，这里发展出的理论属于怀疑论一类。它不同于其它怀疑论的地方在于在此怀疑论理论中被归为扩充的一些句子集在其它理论中并不是这样，反之亦然。取上面一个例子。在赖特尔（Reiter）的常识逻辑的框架之中，这一论证可表示成常识理论<D, *W*>，其中

$$D=\{(p, Mr/r), (q: M\neg r/\neg r), (q: Mp/p)\}，且 W=\{q\}$$

给定赖特尔的扩充定义，这一常识理论有两个扩充：{p, q, ¬r }的演绎闭包和{p, q, r }的演绎闭包，而在这里只考虑第一个。

赖特尔的最初的理论的主要缺点是不回答有关优先的问题。在很多案例中，存在有冲突的规则，一些优先于另一些。上面，我已经试着揭开这一现象背后的机制了。结论有很多和 Delagrande[1988]，Asher&Morreau[1990]相同，它们是建立在 David Lewis[1973]发展起来的条件句语义之上的。

在一方面，这里发展出的理论要比德尔格朗德（Delgrande）和阿谢尔&莫罗（Asher&Morreau）的理论要简单。在检验一个论证的有效性时，三个理论都把目光锁定主体不知道比给定的前提更多的状态。德尔格朗德和阿谢尔&莫罗都试图给出这一状态的直接定义，而在动态框架上它是逐步建立起来的。另一方面，熟悉我提到的那些论文的读者可能会有疑问为什么我不选择选择函数[9]来表示一个主体有关规则的知识。从数学视角来看，这些是比期望框架更简单的对象，而且到此为止，我没有做任何事情来展示让事情像现在这样复杂是真正必要的。

有一个当下理论的更简单版本，其中选择函数被用作一个信息状态的组成部分。在每个案例中，这个简化的版本和现在这个一样好用。实际上，只要我们的案例限制在每个论域至多只考虑一个（非平凡的）的常识的情况。两个版本是一样的。但只要在一个域中有超过一个规则时，差异就产生了。

学生通常是成年人;
学生通常没有工作;
约翰是学生;
约翰有工作;
想必，约翰是成年人。

学生通常是成年人;
学生通常没有工作;
成年人通常有工作;
成年人通常知道怎样开车;
彼得是一个学生;
想必，彼得知道怎样开车。

这些是有时被称为独立律的例子。如果一个对象在一方面是例外的，这不意味着它在其它方面也是例外的。经常，你可以放心它在其它方面是正常的。正如例子所展示的，这一条不仅对偶然例外的对象成立，对非偶然例外的也成立。给定左例的前提，约翰恰巧是规则学生不工作的例外。因此，约翰不是一个正常的学生——至少不是完全正常。然而，没有理由认为规则学生通常是成人不适用。你可仍旧假定约翰是一个成年人。给定赖特尔的理论，你能这么做，给定当下的理论，你也能这么做，但给定德尔格朗德或阿谢尔&莫罗的理论，则不是这样。对右边的例子，一个形式化分析揭示了极佳的彼得——确定尽可能多正常性使用标准

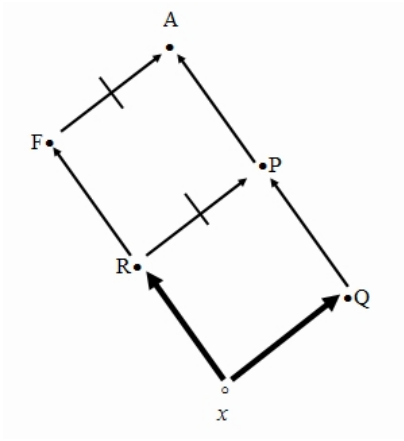
的彼得——是一个对规则成人通常有工作来说为非偶然例外的成年人，但他知道怎样开车。我所知道的其它理论中唯一在这里可以给出些结果的是继承（inheritance）理论，我将在下面说起这一理论。

独立律是有效的主要是因为模式πd 可以不仅仅是 d 是在正常和不正常元素之间的二分法。我相信这一规律包含了常识推理的基本性质。因此，我不能不得出结论，选择函数不是模型化一个主体的规则知识的合适“人选”。

我们的形式化的表达力是有限的。然而，它已经足够丰富到来在一个语义网中表达每一个可表达的东西了。这里给出的理论为对多继承网络提供了一个语义，其中循环路径和复杂谓词是被允许的。它给出一个可判定的非单调的逻辑后承概念，即，有效性$_1$，它与继承理论中的“支持（support）”-关系是可比较的。这可以用作回答可靠性和完全性问题的基础：给定一个合适[10]的网的类推理算法，是不是一个Γ属于支持一个结论φ的类当且仅当它对于从Γ中的规则和事实推出想必φ是有效$_1$的？

对于所有我知道的算法，这个问题的回答是：不是。最接近于答案“是”的算法是Horty&Thomason&Touretzky[1987]给出的。对于我们所讨论过的例子，这一算法和我们当下的理论有相同的输出结果。仍旧，从我们的观点来看，这个算法是不可靠的。如果它可靠，则下一个论证会是有效的。但它不是。

贵格会教徒通常是和平主义者
共和党通常不是和平主义者
和平主义者通常是反军事的
共和党人通常是足球迷
足球迷通常不是反军事的
<u>约翰既是贵格会教徒，又是共和党人</u>
想必，约翰不是反军事的



在我们对于这个论证的形式化中，这是级联含混的例子，有形式

$$q\to p,\ p\to a,\ r\to\neg p,\ r\to f, f\to\neg a,\ q\wedge r/\text{想必}\neg a\text{。}$$

仅仅获知所有这些前提的某人的状态是高度含混的。这里有 4 个极佳对象：{q, r, f}, {q, r, f, p},{q, r, p, a}, {q, r, f, p, a} 。因此，期待 x 是反军事的不有效，x 不是反军事的也不有效。

这里有个例子展示了，从我们的观点来看，霍尔蒂（Horty）等人的算法也不是完全的。否定后件（M.T）的可中断版本是有效的，但表示以下论证的前提的网并不支持结论。

成年人通常有驾照	p→p
<u>约翰没有驾照</u>	<u>¬p</u>
想必约翰不是成年人。	想必¬p

为了看看为什么这个论证是有效的，集合 W={w_0,w_1,w_2,w_3}，其中 w_0=∅, w_1={p}，w_2={q}

且 w_3={p，q}。考虑 **0**[p→q][¬q]。作为‖¬q ‖中的世界，w_0 遵守‖p ‖-常识‖q ‖。但世界 w_1 不是。而且因为对所有扩充‖¬q ‖的域 d，$\mathbf{n}\pi d \subseteq \{w_1\}$，‖p ‖-常识‖q ‖适用于‖¬q ‖之内都不成立。因此，在状态 **0**[p→q][¬q]中的某人会期待现实世界更像 w_0 而不是 w_1。而在 w_0 中，命题‖p ‖不成立。

这启发我们去将上面的论证和接下来这个加以比较：

学生通常是成人	p→q
成人通常不是学生	q→¬p
<u>约翰是一名学生</u>	<u>p</u>
想必，约翰是成人	想必 q

注意，对应于这个论证的继承网是循环的。第一眼看去，这个论证的前件有些含混：根据分离规则（M.P），人们可以推出想必 q，而由否定后件规则（M.T），可推出想必¬q。然而，进一步的检查状态 **0**[p→q][q→¬p][p]揭示了 M.P 是优先于 M.T 的：设 W 如常。重点在于‖q‖-常识‖¬p ‖不适用于‖p‖内，因为 $\mathbf{n}\pi$‖p‖={ w_3}，且{ w_3}⊆‖q ‖~‖¬p ‖。另一方面，‖p ‖-常识‖q ‖适用于‖p‖之内。因此，现实世界会更像 w_3 而不是 w_1，这意味着 **0**[p→q][q→¬p][p] ⊩想必 q。

现在会很清楚了。这篇论文中发展出来的常识理论不仅在解释上和其它理论不一样，在其预测方面也有所不同。没看到你们最喜爱的例子的读者，我将把检验的任务留给你们自己。而这一节，我们将指出一些更具一般性的特点。

首先提醒一下：由我们所用的有效性概念所生成的逻辑不对替换封闭。例如，我们看到，在之前的节中以下论证是有效的。

q→p, p → r , q⊩想必 r (∗)

然而，(∗)仅仅对独立的谓项是有效的——或者至少不知道是不独立的。如果我们用“¬q”替换“r”，我们发现有

q→p, p → ¬q, q⊩想必¬q

作为对第二点的介绍，考虑规则“学生通常是成人”和“成人通常是有工作的”。假设这些是你所知道的所有规则。给定（∗），推出对任意学生 x（对 x 你不知道更多的信息），x 想必有工作是正确的，然而，这不意味着，得出结论“学生通常是有工作的”是正确的。即：

q→p, p → r⊮ q→r (∗∗)

假言三段论是不有效的；我们只有它的一个可废止的版本，(∗)是例证。(∗)的结论是可

废止的。例如，当你得知学生通常不工作时，该结论就被推翻了。而(∗∗)的结论，则不是可废止的。规则是保持的。

有更多这类例子。即使我们经常发现

$$\varphi_1\to\psi_1, \ldots, \varphi_n\to\psi_n, \chi \Vdash_{想必} \theta,$$

然而

$$\varphi_1\to\psi_1, \ldots, \varphi_n\to\psi_n, \nVdash \chi\to \theta。$$

例如，我们知道，M.T 的一个可废止形式是有效的：

$$p\to q, \neg q \Vdash_{想必} \neg p,$$

但假言易位不成立

$$p\to q \nVdash \neg q\to\neg p。$$

我们也有

$$p\to q, p\wedge r \Vdash_{想必} q,$$

但前件加强是不被允许的：

$$p\to q \nVdash (p\wedge r) \to q。$$

我们熟知的蕴涵律像：假言三段论，假言易位和前件加强对常识箭头都不成立。因此，一个很自然的问题是，还有哪些剩下的成立吗？如果箭头→不是一个严格蕴涵，由这些规律的失败所展示的，也许它是严格蕴涵的一个变种？如果它是，则以下给出任意严格蕴涵的变种与经典连结词之间互相作用的完全描述会成立：

条件恒等（CI）[11] : $\varphi\to\psi \Vdash \varphi\to\varphi$
后件合取（CC） : $\varphi\to\psi, \varphi\to\chi \Vdash \varphi\to(\psi\wedge\chi)$
后件弱化（CW） : $\varphi\to\psi \Vdash \varphi\to(\psi\vee\chi)$
用后件加强（ASC）: $\varphi\to\psi, \varphi\to\chi \Vdash (\varphi\wedge\psi)\to\chi$
前件析取（AD） : $\varphi\to\psi, \varphi>\chi \Vdash (\varphi\vee\psi)\to\chi$

但事实上，以上只有前两个是有效的。后面三个是几乎有效的。例如，对任何状态σ，以下成立：

$$\sigma[\varphi\to\psi][\varphi\to\neg(\psi\vee\chi)]=\mathbf{0}$$
$$\sigma[\varphi\to\psi][\varphi\to\chi][(\varphi\wedge\psi)\to\neg\chi]=\mathbf{0}$$
$$\sigma[\varphi\to\chi][\psi\to\chi][(\varphi\vee\psi)\to\neg\chi]=\mathbf{0}$$

对假言三段论，类似的不成立。非常有可能是这样：

$$\sigma[\varphi\to\psi][\psi\to\chi][\varphi\to\neg\chi]\neq\mathbf{0}$$

这有另一个 “几乎有效”的具体（格式）：设Δ是任意规则序列。以下成立：

$$\Delta, \varphi\rightarrow\psi, \varphi \Vdash 想必(\psi\vee\chi);$$
$$\Delta, \varphi\rightarrow\psi, \psi\rightarrow\chi, \varphi\wedge\psi \Vdash 想必\chi;$$
$$\Delta, \varphi\rightarrow\chi, \psi\rightarrow\chi, \varphi\vee\psi \Vdash 想必\chi;$$

这些是 CW, ASC 和 AD 的可废止版本，但它们有一个特殊的性质：它们的结论只能被事实信息来废止。既然

$$\varphi\rightarrow\neg\chi, \varphi\rightarrow\psi, \psi\rightarrow\chi, \varphi \nVdash 想必\chi,$$

因此，同样也和像假言三段论这样的规律有巨大的差异。

我还不能找到一个好的直观解释来解释为什么 ASC 和 AD 不能成立。仅仅对 CW，我有一个论证来展示如果该规律有效的话，有些地方会有问题。事实上，有

$$\varphi\rightarrow\psi \nVdash \varphi\rightarrow(\psi\vee\chi)$$

这里，我重复一下上一节靠近结尾的地方写的东西。如下一个例子所展示的，形为$\varphi\rightarrow(\psi\vee\chi)$的句子在某些方面强于$\varphi\rightarrow\psi$。

—老虎通常有 4 条腿。希瑞 · 坎（Shere Khan）是一只老虎。希瑞 · 坎没有四条腿。所以，想必，希瑞 · 坎有五条腿。

—老虎通常有 4 或 5 条腿。希瑞 · 坎是一只老虎。希瑞 · 坎没有 4 条腿。因此，想必希瑞 · 坎有 5 条腿。

第二个论证是有效的，第一个不是。规则“老虎通常有 4 条或 5 条腿”暗示了当人们遇到不是 4 条腿的老虎时可以期望些什么；规则“老虎通常有 4 条腿”则不。毫无疑问，一个主体也许希望接受后者，而不接受前者。

6. 结束语

这篇文章的目的有两方面：(i)引入更新语义的框架，并且解释哪种语言现象可以在其中被成功的分析；(ii)给出这些现象之一：常识推理的分析。

在更新语义的框架之内，常识推理不是被看成普通句子的特殊推理，而是一类特殊句子的普通类型的推理。从“*x* 是 A”和“A 通常是 B”得出“想必 *x* 是 B”和从“*x* 是 A”，“所有 A 是 B”得出“*x* 是 B”是一样有效的。人们不需要为了获得前者而设定不同的推理模式。在两个案例中，有同样的有效性概念，在处理普通的描述句时，它与单调逻辑中的经典概念是一致的。然而，只要语言丰富到表达常识规则和像“想必”这样的算子，逻辑就变成了非单调的了，因为起始于“想必”的句子是特殊的——它们不保持。

前几节中发展出的有关常识的特殊理论不是在更新语义框架下唯一的可能性。

事实上，人们可能希望在相同的直观想法下，发展出更优雅的形式化结果。我始终认为这些直观想法，其制高点在于一致性标准和适用性标准是可靠的，而且这一理论在预测常识

的变化无常的逻辑行为时比其它理论要好很多是一个事实。

我希望这篇论文中的想法不仅对对常识感兴趣的逻辑学家有所帮助，也同样能对对概称句语义感兴趣的语言学家也有所帮助。我意识到，我这里所提供的充其量只是一个巨型拼图中丢失的一块——没人知道还有多少块还丢着呢，更不用说它们如何组合在一起的问题。我已经给出了一类概称句的逻辑分析，即形为“P 通常是 Q”的句子。不过，不管这样的分析多么有优点，它都没有说这种类型的概称句和其它类型的关系。它没有解释为什么形如

(i) p 通常是 Q

的句子经常与(ii)-(iv)传达相同的信息。

(ii) 这个 P 是 Q
(iii) P 是 Q
(iv) 一个 P 是 Q

它甚至没有解释这样的句子通常等价于：

(v) 通常 P 是 Q。

在 AI（人工智能）背景下，这些句子形式经常互换使用。而且，实际上，从很多实例看来，它们确实有相同的影响。比较以下例子：

(i)′ 老虎通常有 4 条腿
(ii)′ 这只老虎有 4 条腿
(iii)′ 老虎有 4 条腿
(iv)′ 一只老虎有 4 条腿
(v)′ 通常老虎有 4 条腿

但语言学家比逻辑学家更为关注这些句子形式之间的差异。如果句子形式(i)-(v)真的总是等价，我们可以说

(ii)″ 老虎通常灭绝了，

而且(ii)″与(iii)″和有相同的意思。

(ii)″ 这个老虎灭绝了

同时，如果(i)和(ii)真的等价.则

(iii)‴老虎吃人

会蕴涵

(i)‴老虎通常吃人

你又如何考虑下一个句子？

(iv) ′′′一只老虎是可得到的(available)

不管这是什么意思，它与(i)′′′是不等价的。

(i) ′′′′老虎通常是可得到的

这与

(v) ′′′′通常老虎是可得到的

也有很大差异。

这仅仅是围绕概称句的[12]一长列问题中的一个。自从Carlsson[1977]以来，人们很清楚的知道，解决方案部分依赖于合适的谓词子范畴，有些是有关类的，有些是有关个体的，还有其它是有关个体的时间阶段的。但没有理论解释什么时候一个概称句可以获得常识推理，以及这种读法从何而来。这篇论文也没有提供这方面的理论，充其量只是解释了常识读法意味着什么.[13]

参考文献

Asher, N. and M. Morreau: 1990, ‘Commonsense entailment: A modal theory of non-monotonic reasoning’, in J. van Eijck(ed.), *Logic in AI*, Proceedings of JELIA ′90，Springer Lecture Notes in Computer Science 478, 1-30.

Beaver, D.: 1991, ‘The kinematics of presupposition’, *Proceedings of the 8th A′dam Colloquium.*

Carson, G.: 1977, *Reference to Kinds in English*, Ph.D dissertation, University of Massachusetts, Amherst.

Delgrande, J.: 1988, ‘An approach to default reasoning based on a first-order conditional logic: Revised report’, *Artificial Intelligence* **36**, 63-90.

Gärdenfors, P.: 1984, ‘The dynamics of belief as a basis for logic’, *British Journal for the Phiosophy of Science* **35**, 1-10.

Groenendijk, J. and M. Stokhof: 1991, ‘Dynamic predicate logic’, *Linguistics and Philosophy* 14, 39-101.

Groeneveld, W.: 1989, ‘Dynamic Semantics and Circular Propositions’, *Journal of Philosophical Logic* **23**, 267-306.

Heim, I. R.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. Dissertation, University of Massachusetts, Amherst.

Horty, J., R. Thomason, and D. Touretzky: 1987, *A Skeptical Theory of Inheritance in Non-Monotonic Nets*, report CMU-CS-87-175, Carnegie Mellon University, iii+52pp.

Kamp, J. A. W.: 1981, ‘A theory of truth and semantic representation’, in J. Groenendijk, T. M. V. Janssen, and M. Stokhof (eds.), *Formal Methods in the Study of Language*, Mathematical Centre Tracts 135, Amsterdam, 277-322.

Krifka, M.: 1987, ‘*An Outline of Genericity*’, SNS-Bericht 87-25, Seminar für Natürlich-Sprachiliche Systeme, Universität Tübingen.

Lewis, D.: 1973, *Counterfactuals*, Basil Blackwell, Oxford.

Reiter, R.: 1980, ‘A logic for default reasoning’, in *Aritficial Intelligence* **13**, 81-132.

Stalnaker, R.: 1974, ‘Pragmatic Presuppositions’, in Munitz, M. and P. Unger(eds.), *Semantics and Philosophy*, University Press, New York, 197-213.

Van Benthem, J. F. A. K.: 1991, *Language in Action*, Elsevier Science Publishers(North-Holland), Amsterdam.

Zeevat, H: 1992, ‘Presupposition and accommodation in update semantics’, *Journal of Semantics* **9**, pp 379-412.

---

[1]含义（meaning）这个概念以最近形式语义学中的很多工作为基础，其起源可以追溯到罗伯特•斯托内克尔（Robert Stalnaker）有关预设和断言的工作（例如，Stalnaker[1974]）。它在汉斯•坎普（Hans Kamp）和艾琳•海姆（Irene Heim）有关照应（anaphora）的工作以及彼得•贵登福斯（Peter Gärdenfors）在信念的动态方面的工作中进一步成型（见例 Kamp[1981], Heim[1982], Gärdenfors[1984]）。最直接的灵感来自吉荣•胡能迪克（Jeroen Groenendijk）和马丁•斯托克霍夫（Martin Stokhof）在动态谓词逻辑中的工作（见 Groenendijk, J. and M. Stokhof [1991]）。

[2] 详见 Beaver[1991]和 Zeevat[1992]。

[3] 在读者可能倾向于使用“信念”和“信念状态”的地方，我使用了短语“知识”和“知识状态”。实际上，我想用信息状态σ表示两者之间的东西：如果σ是一个给定主体的状态，它应该表示的是该主体认为是他/她的知识的东西。主体仅当做信念的东西不算。但非常有可能主体认为已知的东西事实上是假的。

[4] 在当前的系统内形式化这一例子是不可能的。因为一个信息状态$<\varepsilon, s>$中的集合 s 刻画该主体“现实”情境下的知识。如果在状态$<\varepsilon, f>$下会有更具一般性的结果，其中ε是 W 上的模式（如前），f 是一个函数，给每个时间点 t 指派 W 的一个子集 $f(t)$，表示该主体在时间 t 的知识。这样做，我们也可以形式化处理主体有关明天天气的期望。

[5] 在试图抓住一个信息状态的定义的过程中，读者可能会疑惑为什么状态$<\varepsilon, s>$中的模式ε是 W 上的模式，而不是 s 上的，$<\varepsilon\cap(s\times s)>$能不能做到现在$<\varepsilon, s>$所做的？这一问题的答案是“不”：如果在那样的定义下，规则们将不再保持了。

[6] 这里，为了集中注意力。我把语言限制在必然算子只能作为一个句子最外层的算子出现的情况。把这一理论扩展到不仅缺省规则，且严格规则也能被我们主体所理解并不难。这里是基本想法：一个状态σ是$<\varepsilon, s>$如前，只是ε现在是 W 的子集 V 上的模式，而不是之前的 W 自己上的模式。如下一个更新从句所展示的，V 由主体所知的严格规则所决定：一个世界 w 是 V 中的一个元素，仅当，每个该主体认为是必然的命题在 w 中都成立。

- 如果 $\mathbf{n}\varepsilon\cap\|\varphi\|=\emptyset$或 $s\cap\|\varphi\|=\emptyset$，$\sigma$[必然$\varphi$]=**1**；
- 否则，σ[必然φ]=$<\varepsilon|(V\cap\|\psi\|)s\cap\|\psi\|>$。

[7] 我不能证明以下是不可能的：存在一个更新系统，其中以下成立：

(i) 通常 p$\Vdash$通常(p$\vee$q)

(ii) 通常 p, $\neg$p$\nVdash$想必 q

(iii) 通常(p$\vee$q), $\neg$p$\Vdash$想必 q

然而，如果你想要这样一个系统，就不得不放弃矢列切割（假设(i)成立，给定矢列切割和(ii)，有通常 p，通常(p∨q)，¬p⊮想必 q，但这几乎和没有(iii)一样糟糕）。

[8] 如果把“如果 q，则通常¬p”和“q⊃通常¬p”等同起来也不行。该句充其量带着主体进入一个状态，其中他们相信或者 q 在现实世界中偶然（刚好）假，或者 p 通常不真。（注：官方的 q⊃通常¬p 不是$\mathcal{L}_2^A$中的句子）

[9]一个选择函数是函数 *f*，给每个 W 的子集 *d* 指派一个 *d* 的子集 *f*(*d*)，直观上，f(*d*)包含 *d* 的正常元素。

[10]这里，“合适”意味着“每件能在网络语言中说的事情，在语言$\mathcal{L}_3^A$中也能说。”我以下将不会严格的区分这两者。

[11]在很多条件句逻辑中 CI 没有限制的成立：$\Vdash\varphi\rightarrow\varphi$。我在这里不得不加个说明：$\varphi\rightarrow\varphi$只对非-矛盾的$\varphi$是有效的。（在谬环境下没有事情是正常的）

[12] Krifka[1987]中有超乎想象的概述。

[13]本文有很多草稿。第一个发表在 ESPRIT. Basic Research Action3175,Report 2.5.A DYANA, Centre for Cognitive Science, Edinburgh, 1990。