更新模型下一类推理难题的形式化解释[1]

郭佳宏
（中山大学哲学系， 广东广州 510275）

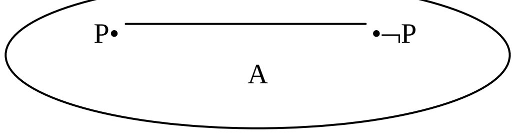
摘要：近几十年来，逻辑学家们的兴趣发生了较大的变化，他们已经把认知动作和过程作为中心问题来研究，即所谓的“动态转向”。其中主要的流派有动态认知逻辑、更新逻辑、信念修正和分支时间等。van Benthem 等[vBen01]，[vBen03]研究了“公共宣称”的转换模式；Baltag 等[BMS98]，[BMS03]提出了更新积模型，并且首次引入了状态模型之间“更新关系”及更新模态算子。本文并不打算对相关的理论作详细的介绍，而是试图用其中的“转换模式”和认知更新模型对日常生活中某些涉及知识变化的直观推理难题作一番精确的分析，从而试图表明这样的模型还是能够比较好地刻画生活中某个局部知识更新的直观意义的。其间，那些涉及“不成功更新”的推理难题得到了严格地解决，并且某些隐含的问题被挖掘出来，比如冗余性和时态等。在上述动态模型下，这类推理难题的结构被清晰地展现出来。

关键词：公共宣称 推理难题 不成功更新 更新模型 更新逻辑
中图分类号：**B81**　　**文献标识码**：**A**

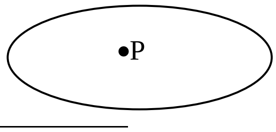
1 引言

现代逻辑的动态转向使得静态模型和动态描述成为了两个并列的成分。在认知逻辑领域，静态模型是多主体信息模型，动态部分则是分析信息模型如何在宣称或某些动作下受到激发，从而发生自然更新、转换等过程。静态模型早就出现于认知逻辑中[FHMV95]，在那里，可能世界表示跟主体相关的她们认为可能的状态；每个主体对应可能世界集上一个二元不确定关系（uncertain relations），直观表示某主体无法区分对应的两个世界。不确定关系是等价关系。而动态部分的形式体现于原静态模型的变化，导致模型不断变化的因素是认知动作及原模型的结构。这里有一个非常简单的例子[vBen03]可以对此作出直观说明（复杂的例子有 Fagin 等[FHMV95]的“泥孩难题”， van Benthem [vBen01]的“猜牌游戏”等）。

例 1：A 问 B：“这条路通向广州吗？”B 回答：“是的”。我们记“这条路通向广州”为 P，并且根据一般情况假定：A 不知道 P；她相信 B 能对 P 有所断定（事实上 B 知道 P 并且 P）。于是能够表达 A 和 B 的认知状况的可能世界模型如下：



这里，P 和¬P 对应两个可能世界（记为 w_1 和 w_2），分别表示 P 在其中为真和为假，其他原子命题在这两个世界中的取值无关紧要。w_1 对应的为现实世界。A 无法区分这两个世界，而 B 可以区分。这两个世界对 A 和 B 来说是自反的，严格地说还应在图上标出自反关系。在此模型中，我们有 $w_1 \models K_B P$ 但 $w_1 \models \neg K_A P$。虽然 A 不知道 P，但她知道或者 B 知道 P 或者 B 知道¬P，即有 $K_A (K_B P \vee K_B \neg P)$。当 B 作出肯定回答后，对 A 而言，她就可以区分原来不能区分的两个可能世界并且知道 P。于是原来的模型变成：

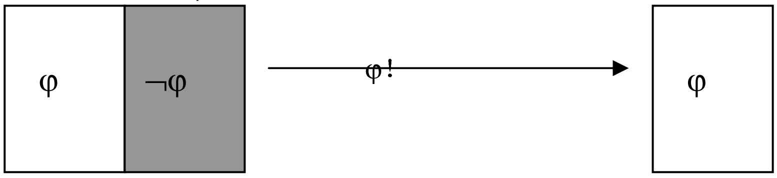


[1] 本研究受中山大学优秀博士论文培育项目、中山大学凯思奖学金资助

收稿日期：2005 年 9 月

不能满足 P 的 w_2 被删除了，相应的不确定关系也消失。此时，A 和 B 都知道 P，并且 B 知道 A 知道 P，A 知道 B 知道 A 知道 P，…。我们称 P 为 A 和 B 的公共知识，记作 $C_{\{A, B\}}P$。

B 的回答对 A 和 B 这个群体来说是一个公共宣称（public announcement）。van Benthem [vBen03]给出了有关公共宣称的简单模型转换模式：关于φ的断定的公共宣称φ！消除当前模型中所有那些不能满足φ的世界。简单图示如下：



上面的解释可以简单地处理认知动作（epistemic actions），但即使这样，我们并没有把认知动作放到核心的位置，按照 van Benthem [vBen03]的说法，它们还是“二等公民”。认知逻辑的语言并没有把它们包括在内，所以在这一范围内没有相应的演算系统来直接地对它们进行推理。

逻辑学家们在寻找新的工具。能够描述命题和动作共存的一种早先的逻辑是“动态逻辑”[KHT84], [KHT00]，其核心思想是每一个动作都对应模态算子，即[a]φ是语言的一部分。a 和φ分别表示动作和命题，[a]φ直观表示“经过每一个成功执行动作 a 后φ成立”。a 也可以是交流动作，如我们前面提到的公共宣称，这样就可以用逻辑语言直接表达它。$[A!]K_j\varphi$表示经过有关 A 的真实公共宣称后，主体 j 知道φ。在此基础上建构的逻辑称之为动态认知逻辑，实际上是动态逻辑和认知逻辑结合的产物。于是复杂的动态认知规律可以得以表达，比如有效式$[A!]\ K_j\varphi \leftrightarrow (A\rightarrow\ K_j(A\rightarrow[A!]\ \varphi)$把公共宣称后获得的知识跟主体原先知道的联系起来了。

前面提到的公共宣称是人类交流中很重要的一类动作，但是并非所有交流都是公共的，有许多形式的信息转换是在完全隐蔽的或者半透明的情况下进行的，甚至还会涉及到主体的作弊、相互欺骗等行为。相应地由这类动作引起的认知模型更新就要比删除可能世界的公共宣称复杂得多。已有不少学者对此类更新动作的一般情况进行了研究，如 van Ditmarsch [vDit00]，Baltag et al.[BMS03]，van Benthem & Liu [vBL04]等。我们把这类工作统称为更新逻辑（update logic）。

我们从文章的第二部分开始介绍一些经典的认知推理难题，如扑克牌难题、生日难题及数字难题，这是一类很有代表性的涉及知识变化的推理，我们将仔细分析其推理过程，在文章的第三部分将通过构造认知推理模型，进行分阶段状态模型的动态分析。通过动态分析，此类推理涉及的时态问题、冗余性问题将清晰呈现出来，文中作者对冗余性问题进行了详细的讨论并给出了消除冗余的方法。在文章的最后部分作者给出了此类推理难题在更新积模型下的更一般解释。

2 认知推理难题

有一类经典的认知推理难题，主要涉及在半透明的前提下“从公共宣称不知道到导致知道”这样的不成功更新（unsuccessful update, vDit00）动作。它们看起来比较复杂，以至某些国际大公司采用这样的难题来测试应聘者的推理能力。具体的有以下几个代表：

例 2：假设主体 a, b 和 c 有足够的推理能力并且他们真实地说话，d 告知大家抽屉里有 16 张牌，分别是红桃 A, Q, 4；黑桃 J, 8, 4, 2, 7, 3；草花 K, Q, 5, 4, 6；方块 A, 5。

现在 d（不同于 a, b, c）从 16 张牌中拿出一张，告诉 a 牌的点数，告诉 b 牌的花色 （a 和 b 被告知了点数和花色这件事成了公共知识），然后问 a 和 b 是否能推出这张牌是什么。接着，

c 听到如下的谈话：
a：“我不知道这张牌。”
b：“我早就知道你不知道这张牌。”
a：“我现在知道了。”
b：“我现在也知道了。”
c 通过这几句话就可以推出这张牌到底是什么。请问这张是什么牌？

例 3：张老师的生日是 M 月 N 日，是下列 8 种可能性中的一种：3 月 4 日，3 月 5 日，3 月 8 日；6 月 4 日，6 月 7 日；9 月 1 日，9 月 5 日；12 月 1 日，12 月 2 日，12 月 8 日。他分别告诉小明和小红（他俩的推理能力足够强并且说真话）M 值和 N 值（同样小明被告知 M 值和小红被告知 N 值成了公共知识）。接着小明和小红的对话如下：
小红：我不知道张老师的生日。
小明：在你说这句话之前，我就可以断定，“如果我不知道，那么你也不知道”。
小红：现在我知道了。
小明：现在我也知道了。
根据上述对话，请问张老师的生日是什么？

例 4：两自然数 m 和 n 满足条件 $2 \leq m \leq n \leq 99$。在 S 先生和 P 先生都在场的情况下，S 先生被告知两数的和，P 先生被告知两数的乘积。P 先生和 S 先生的推理能力足够强。P 先生和 S 先生的对话如下：
P：我不知道这两数字。
S：我早就知道你不知道，不过我也不知道。
P：现在我知道这两数字是什么了。
S：现在我也知道了。
根据上述对话，m 和 n 究竟是多少？[McC78-81]

尽管上述的例子还有缺陷，但一般情况下我们都能得到相同的结果，而且自然推理的过程也大致相同。比如对于例 2，c 可能作了如下推理：最初，a 知道点数但是并不能断定这张牌，因此这张牌一定不是 J, 8, 2, 7, 3, K, 6 。如果是这 7 张中的任何一张的话，a 有足够的推理能力，则可以直接断定。所以，牌的点数只可能是 4 、5 、A 或 Q。再分析 b 的断言，b 从牌的颜色在 a 宣称之前做出了下述推论：a 不知道这张牌。这意味着牌的颜色不能是黑桃或草花，如果是黑桃或草花中的一张，b 没有足够的理由在 a 说他不知道这张牌之前断定 a 不知道。因此，c 能够推出这张牌是红桃或方块。同样的信息能够传递给 a 和 b，因为他们的推理能力足够强。进一步，当 a 听到 b 的话后，a 可以对这张牌作出断定，则这张牌一定不会是 A，假设这张牌是 A，a 仅知道点数和颜色范围（红桃或方块），他并不能推出这张牌到底是什么——这张牌可能是红桃 A，也可能是方块 A。至此对 c 来说，只剩下三种可能了：红桃 Q，红桃 4 或方块 5。相似的，再由 b 的话，c 能够推出这张牌一定是方块 5。推理如下：b 只知道牌的颜色，如果是红桃的话，b 不能断定到底是红桃 Q 还是红桃 4，而现在 b 可以断定牌是什么，则说明这张牌一定不是红桃，所以这张牌是方块 5。

例 3 的情况稍微有所不同，小明的断定用了条件句。但实质上跟例 2 是一样的，因为根据初始情况，这个条件句的前件是成立的；即我们可以得到后件“我早知道你不知道”。在后面的形式化分析中，我们将说明例 3 跟例 2 一样，对话中的第一句宣称是多余的。然后我们将构造一个例 2 的变种，使得其中涉及的对话里每一句都对达到正确目标有信息更新作用。

例 4 的情况较为复杂，初始的可能性太多，而且还涉及到自然数理论的基础知识。不过认知推理过程的结构也是类似的。当 P 先生宣称不知道后，我们首先可以排除这两数字都是质数的情况，因为如果它们都是质数，P 先生就知道这两数了。当 S 宣布后，我们可以排除这两数之和为偶数的情况，因为上述涉及的 200 以内（根据歌德巴赫猜想获得的观测结果）大于 2 的偶数都可以表示成两个质数之和，而这种情况下 P 就能知道了，这样 S 就不能那样宣称他知道 P 不知道了；同理，那些两数之和可以表示成质数加 2 的奇数也需要排除掉；另外我们还可以排除那些两数之和大于 53 的情况，因为任何一个大于 53 的奇数可以表示成一个大于 50 的质数和一个偶数之和。考察乘积与此质数和偶数之积相同的另外两数，不难发现其中有一个一定超过 100，矛盾于题设条件。那么也说明在这种情况下，合理的可能性乘积只有一种，P 于是就知道了，S 不能宣称他早知道 P 不知道。既然事实上 S 宣称了他早知道 P 不知道，那么这种情况也同样被排除了。这样在 S 宣称后，两数之和的可能性就只剩下以下几种了：11，17，23，27，29，35，37，41，47，51，53。我们列出所有上述两数之和各种可能性中的所有两数之积的可能性，分别如下：

11（18,24,28,**30**）

17（**30**,**42**,52,**60**,**66**,**70**,**72**）

23（**42**,**60**,76,**90**,**102**,112,**120**,**126**,130,**132**）

27（50,**72**,92,110,**126**,140,152,162,170,176,182）

29（54,78,100,**120**,138,154,168,**180**,190,198,204,208,210）

35（**66**,96,124,**150**,174,**196**,216,234,250,264,276,**286**,294,**300**,304,306）

37（**70**,**102**,**132**,160,186,**210**,232,252,**270**,**286**,**300**,**312**,322,**330**,336,340,**342**）

41（78,114,148,**180**,**210**,238,264,288,310,**330**,348,364,378,390,400,408,414,418,420）

47（**90**,**132**,172,**210**,246,280,**312**,**342**,370,396,420,442,462,480,496,510,522,532,**540**,546,550,552）

51（98,144,188,230,**270**,308,344,378,410,440,468,494,518,**540**,560,578,594,608,620,**630**,638,644,648,650）

53

（**102**,**150**,**196**,240,282,322,360,396,430,462,492,520,546,570,592,612,**630**,646,660,672,682,690,696,700,702）

不难发现并证明上述各行数中的增量成等差递减数列，等差为 2，所以就不难把上述的所有数值列出来。当然也可以更简洁地把上述各行数表示数列形式，全部列出来的好处是可以直观地发现重复出现的数字。接着，知道两数之积的 P 宣称他知道了，于是上述数值中那些重复出现的数字必须被排除掉，否则 P 就不会宣称了。再接下去，知道两数之和的 S 也宣称知道了。我们考察去掉那些重复数字后的情形。在两数之和的所有可能性中，17 这一行只剩下一种乘积的可能，即 52，其他行中都出现大于两种的可能性。既然 S 宣称他也知道了，那么那些乘积的可能性还大于两种的两数之和的可能性都要被排除掉。所以最后只剩下 17 这种可能，而对应的两数之积为 52。这样我们可以得到这两数是 4 和 13。

有关此例的一阶逻辑形式化请参阅[McC78-81]，那里的情形看起来较为复杂。当然如果把它编成程序，复杂的计算可由计算机代替。

3 认知更新模型下的结构

3.1 分阶段状态模型动态分析

我们主要针对例 2（例 3 和例 4 的本质过程类似）。为了较简单的描述这个过程，我们把牌的数量减少到 10 张，分别为：红桃 A, 4, Q; 黑桃 8, 4; 草花 K, 5, Q; 方块 A, 5。不难发现，推理过程和上述 16 张牌的推理过程在本质上是一样的。首先引入涉及主体知识的语言 L：

定义 1： L 为在经典的命题逻辑语言 L 上增加算子 K_i 后的扩张语言，其中 K_i（i∈*A*，*A* 为主体集{a, b}）为知道算子。

定义 2：K_i 的对偶算子记作<i>，$\langle i\rangle A =_{df} \neg K_i \neg A$（A 为 L 中的任意公式）。

经典的知道逻辑极小系统为 S5。接下来我们用这 10 张牌作为域建立状态模型，给出有关模型的相关定义。

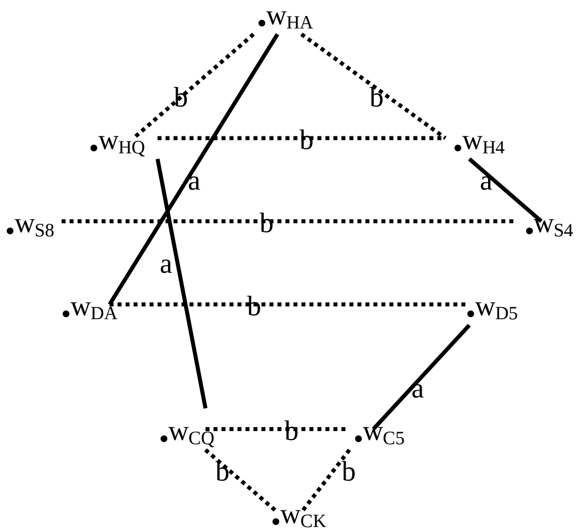
定义 3：状态模型为四元组 $M = \langle W, R_a, R_b, [\ ]_M\rangle$，其中，

$W= \{w_{HA}, w_{HQ}, w_{H4}, w_{S8}, w_{S4}, w_{DA}, w_{D5}, w_{CQ}, w_{C5}, w_{CK}\}$，为状态集，直观表示 10 种可能的情况；

$R_a=\{(w_{Xi}, w_{Yj})|\ i=j: X, Y\in\{H, S, C, D\}, i, j\in\{A, 4, Q, 8, K, 5\}$，为二元关系，直观表示 a 不能区分的状态序对；$R_b=\{(w_{Xi}, w_{Yj})|\ X=Y: X, Y\in\{H, S, C, D\}, i, j\in\{A, 4, Q, 8, K, 5\}$，为二元关系，直观表示 b 不能区分的状态序对；$[\ ]_M$ 表示从原子句子到 W 上的映射。

定义 4：我们规定有且只有 10 个原子句子，分别表示“这张牌是红桃 A”，“这张牌是红桃 4”，…分别记作 HA, HQ, H4, S8, S4, DA, D5, CQ, C5, CK, 称此集为 AtSen。同时我们规定，对任意的 $p\in$AtSen，$[p]_M=\{w_p\}$。

这个模型其实已经是在描述 a 和 b 被告知点数和花色后的情况。严格来说，我们应该从初始状态出发，不过由于此模型已经可以直观把握，为了简化分析过程，我们就把它作为初始模型，记作 M_1。显然 M_1=M。以下是 M 的图形表达：



M_1: 在我们分别告诉 a 和 b 牌的点数和花色后（自反关系省略，下同）

接下来考虑可满足关系：

定义 5：$w\models_M p\ (w\in W)$，如果 $w\in[p]_M$ 对于任意的 $p\in$AtSen（如果在特定语境下不至混淆，我们将省略 M 记号）。根据经典的 Kripke 关系语义，[ ]的定义域可以扩展到语言 L 上，各算子情况分别如下：$[\neg A]=W-[A]$，$[A\wedge B]=[A]\cap[B]$，$[K_iA]=\{w\in W|$如果 wR_iu，那么 $u\in[A]\}$，其中 A，B 为 L 中的公式。归纳于公式的复杂性，不难证明：

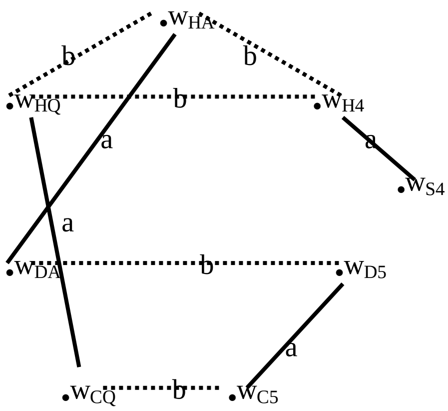
命题 1：对于任意 L 中的公式 A，$[A]\subseteq W$。（证明略）

于是，我们可以给出语义后承的定义：

定义 6：对于任意 L 中公式 A，$w\models A(w\in W)$，如果 $w\in[A]$。

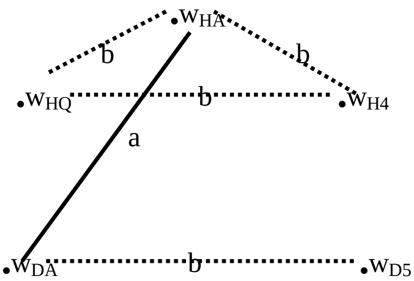
根据上述初始模型和 van Benthem [vBen 03]的公共宣称模式，我们现在开始分析整个更新过程。

阶段 1： 在 a 宣称“我不知道这张牌是什么”之后，表明在新模型中对任意 $w\in W$ 和 $p\in$AtSen，$w\models\neg K_ap$。我们删除 M_1 中不满足条件的状态，于是原模型被更新为 M_2。状态 w_{S8} 和 w_{CK} 被删除，因为根据初始定义，我们有 $w_{S8}\models_{M1}S8$，$w_{CK}\models_{M1}CK$，w_{S8} 和 w_{CK} 都不与其他的（自己除外）状态具有 R_a 关系，于是有 $w_{S8}\models_{M1}K_aS8$，$w_{CK}\models_{M1}K_aCK$。



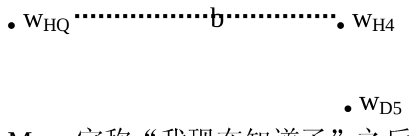
M_2: a 宣称“我不知道这张牌是什么”之后

阶段 2：接下来 b 宣称“我早知道你不知道”。这里涉及过去时，意味着我们必须对 a 宣称之前的模型进行考虑，不能直接对 M_2 更新。在 M_2 之前只有 M_1，所以在 b 宣称之后，我们再次考查 M_1。由于 b 的宣称，我们知道 M_1 中不能满足“b 知道 a 不知道”的所有状态将被删除。而在 M_1 中有关黑桃和草花的相应状态上，即 $w_i\in\{w_{S4}$，w_{S8}，w_{CQ}，w_{C5}，$w_{CK}\}$，有 $w_i\vDash\neg K_b\neg K_aj$（$\exists j\in\{S8, CK\}$）。例如在 w_{S4} 中，$w_{S4}\vDash\neg K_b\neg K_aS8$。因为在 M_1，$w_{S8}\vDash K_aS8$，而 b 不能区分 S8 和 S4，在 S4 她认为 a 知道 S8 是可能的，所以有 $w_{S4}\vDash\langle b\rangle K_aS8$，即 $w_{S4}\vDash\neg K_b\neg K_aS8$。其他的情况类似。但既然 b 宣称她知道 a 不知道这张牌，表明在新模型中，对任意的状态 w，有 $w\vDash K_b\neg K_ap$（$\forall p\in AtSen$）。这就意味着所有草花和黑桃相应的状态必须被排除，因此模型 M_1 被更新为 M_3：



M_3: b 宣称“我知道你不知道这张牌”之后

阶段 3：在新的情形下，a 宣称她知道了这张牌是什么。而在 M_3，对任意的 $p\in\{HA, HQ, H4, DA, D5\}$，有 $w_{HA}\vDash\neg K_ap$，以及 $w_{DA}\vDash\neg K_ap$。因此，我们可以删除 w_{HA} 和 w_{DA} 这两个状态，于是 M_3 更新为 M_4：



M_4: a 宣称“我现在知道了”之后

阶段 4：在此之后，b 宣称她也知道了这张牌是什么。但在 M_4 有，$w_{HQ}\vDash\neg K_bHQ$ 和 $w_{H4}\vDash\neg K_bH4$，因此 w_{HQ} 和 w_{H4} 也被删除。于是得到新的模型 M_5，在此模型中只有一个状态 w_{D5}。而在此模型下，显然有 $w_{D5}\vDash K_aD5$，$w_{D5}\vDash K_bD5$。由于 c 能分析出上述的整个过程，这就意味着她也知道了这张牌是 D5。

M_5: b 宣称“我现在也知道了”之后

3.2 冗余问题及其消除

在例 2 中（例 3 和例 4 的情况类似），不难发现，a 的第一个宣称对 c 最后得出结论来

说是冗余的。对于 c，仅有 b 的宣称和在 a 的宣称之后再有 b 的宣称，更新的结果是相同的。所以我们可以抛弃 a 的宣称，于是 a 和 b 的对话可以缩减为：

b："我知道你不知道这张牌是什么。"

a："我现在知道了。"

b："我现在也知道了。"

这样对话中就没有涉及过去时的宣称，处理起来更为容易。进一步，我们用玩家 1 和玩家 2 来代替一般的主体双方，不难得到以下更为一般的结论：

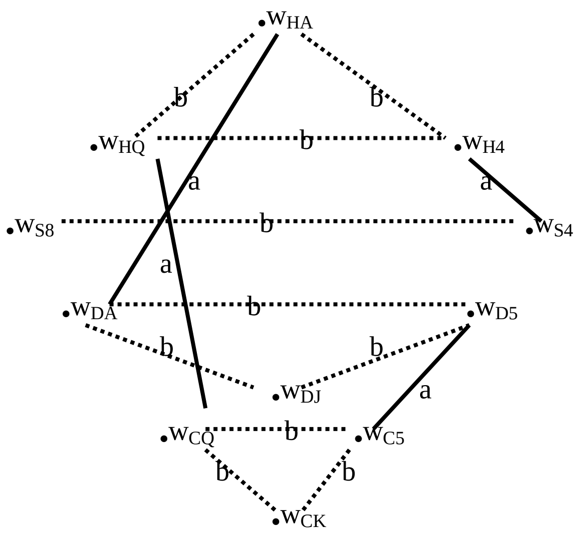
命题 2：玩家 1 宣称她不知道 p，接着玩家 2 宣称她早知道玩家 1 不知道 p，玩家 1 的宣称对类似 c 的听者是冗余的。

证明：对任意的模型 M_1，在玩家 1 宣称$\neg K_1p$之后，对任意 $w_i \in M_1$，使得 $w_i \models K_1p$ 被删除，M_1 被更新为 M_2。接着，玩家 2 宣称 $K_2{}^E{}_1\neg K_1p$ ($K_2{}^E{}_1$表示玩家 2 在玩 1 宣称之前就知道)。这意味着对任意 $w_j \in M_1$，使得 $w_j \models \neg K_2\neg K_1$ 被删除。由于 R_2 是等价关系，$w_i \models K_1p$ 蕴涵 $w_i \models \neg K_2\neg K_1p$。因此，满足 K_1p 的 w_i 也在玩家 2 的宣称后被删除。这意味着在第一次更新过程中所删除的任意 w_i 均包含在第二次更新过程所删除的状态中。 ┤

接下来，我们对例 2 稍作修改，使得其中的每一句公共宣称对于 c 最后结论的得出都不是冗余的。为简单起见，我们同样只考虑 10 张牌。除了下述两点外，其他与例 2 相同：1）d 突然在 a 说话之前对大家说："对不起，我漏了说一张牌，方块 J；实际上牌的总数量是 11 张。J 这张牌也可能是其中的一张"，记此宣称为 B!（宣称的内容为 B，整个宣称动作为 B!）。2）在 a 说"我不知道这张牌是什么"之后，b 说"在 d 宣称 B 之前，我知道你不知道这张牌是什么"。

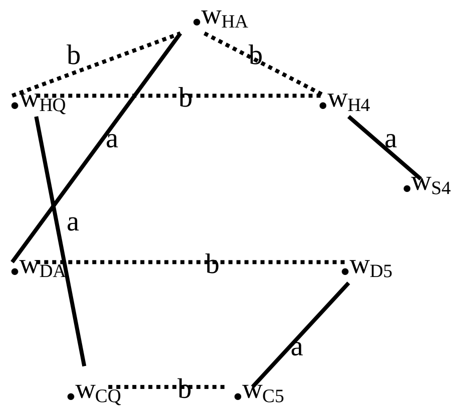
在这个新的例子中，易见 a 最初的宣称对更新后的最后结果不是冗余的。我们分析 c 的更新过程。假定涉及的主体都有足够强的记忆能力：每一个主体在每一次成功更新后，都可以把相关信息记录下来。

阶段 1'：开始阶段的模型同例 2 中的 M_1，但 d 的宣称使得它变得复杂。



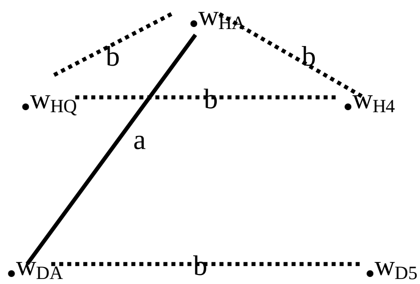
(M'$_2$)

阶段 2’：a 宣称后，把 M'_2 更新为：



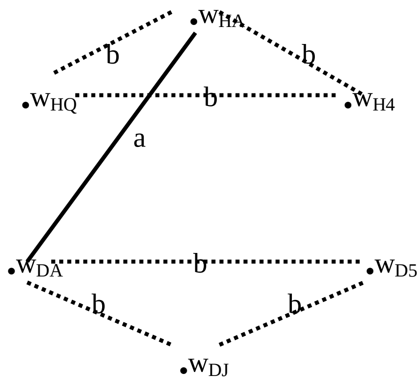
(M'_3)

阶段 3’：然后根据 b 的宣称中的时间关系，首先回到 M_1。接着采用宣称内容“知道 a 不知道”，把模型 M_1 更新为：



(M'_4)

但实际上，由于 d 过去宣称过 B，c 应该把这个信息加到模型 M'_4 中，于是模型变为：



(M'_5)

阶段 4’：而在 d 宣称了之后，a 说过“她不知道”，原来的有关从 M'_2 到 M'_3 的更新信息被保留下来，c 能应用这个更新信息去更新 M'_5，于是模型重新变为 M'_4。可见，如果没有 a 的宣称，c 在 M'_5 之后就无法进行下去了。

接下来的更新同例 2 中的分析，最后的结论也相同。

3.3 更新积模型下的更一般解释

接下来，根据[BMS03]和例 2 中的 M，我们将构造相应的更新积模型：

定义 7：更新积模型 $\mathcal{M}$= <W⊗Σ, R_a, R_b, $[\ ]_{\mathcal{M}}$>，其中 W∈M。

首先需要说明的是简单动作模型**Σ**=<Σ，R_a，R_b，pre>。其中

Σ表示简单动作集。在例 2 中，四个阶段的动作集分别是Σ_1={σ1}，Σ_2={σ2}，Σ_3={σ3}，Σ_4={σ4}，其中对应的动作分别是“a 宣称不知道”，“b 宣称早就知道 a 不知道”，“a 宣称知道”和“b 宣称知道”。

R_b 和 R_a 分别表示 a 和 b 无法区分的某两个动作组成的序对集。例 2 中她们都对$\Sigma_i(1\le i\le 4)$中的每个动作与自己本身无法区分。

pre: Σ→Φ（不涉及动作的公式集），表示动作执行的先决条件。在例 2 中，$pre(\sigma1)=\neg K_aHA\wedge\neg K_aH4\wedge\ldots\wedge\neg K_aCK$, $pre(\sigma2)=K_b(\neg K_aHA\wedge\neg K_aH4\wedge\ldots\wedge\neg K_aCK)$, $pre(\sigma3)=K_aHA\vee K_aH4\vee\ldots\vee K_aCK$, $pre(\sigma4)=K_bHA\vee K_bH4\vee\ldots\vee K_bCK$。

于是我们可以定义更新积模型中的各个元素：$W\otimes\Sigma=\{(w,\sigma)\in W\times\Sigma: w\in[pre(\sigma)]_M\}$; $(w,\sigma)R_a(w',\sigma')$ iff wR_aw'并且$\sigma R_a\sigma'$, $(w,\sigma)R_b(w',\sigma')$ iff wR_bw'并且$\sigma R_b\sigma'$; $[p]_{\mathcal{M}}=\{(w,\sigma)\in W\times\Sigma: w\in[p]_M\}$。

定义 8：更新语言 L’为 L 基础上增加模态算子[]和动作符号后的扩张语言，有关动作的符号集记为 **Term**，所有的公式集记为 **Form**。其中：

p_i∈**Form**；

如果 A, B∈**Form,** 那么¬A, K_iA, A∧B∈**Form;**

σ∈**Term**（σ为简单动作符号）；

如果β,γ∈**Term**，那么β;γ，β+γ，β*∈**Term**（有关;, +, *的定义参阅[KHT00]）；

如果β∈**Term** 并且 A∈**Form**，那么[β]A∈**Form**；

如果 A∈**Form**，那么 A!∈**Term** (直观表示公共宣称了 A)。

根据 M 的性质和上述定义，我们有：

命题 3：$M\models[\sigma1][\sigma2][\sigma3][\sigma4](K_aD5\wedge K_bD5)$ 或者等价地记为

$M\models[(\neg K_aHA\wedge\neg K_aH4\wedge\ldots\wedge\neg K_aCK)!][(K_b(\neg K_aHA\wedge\neg K_aH4\wedge\ldots\wedge\neg K_aCK))!][(K_aHA\vee K_aH4\vee\ldots\vee K_aCK)!][(K_bHA\vee K_bH4\vee\ldots\vee K_bCK)!](K_aD5\wedge K_bD5)$

证明：分析更新过程。由于每一个阶段的动作都是公共宣称，对 a 和 b 来说在相应的每一个阶段只有一个动作，并且她们对这个动作与自己本身无法区分。于是整个更新过程可以看作是 4 个动作按照先后顺序更新总和，即 $\mathcal{M}1$= <W_0⊗{σ1}, R_a, R_b, $[\ ]_{\mathcal{M}1}$>，$\mathcal{M}2$= <W_0⊗{σ2}, R_a, R_b, $[\ ]_{\mathcal{M}2}$>，$\mathcal{M}3$= <W_2⊗{σ3}, R_a, R_b, $[\ ]_{\mathcal{M}3}$>，$\mathcal{M}4$= <W_3⊗{σ4}, R_a, R_b, $[\ ]_{\mathcal{M}4}$>。其中 W_1= W_0⊗{σ1}，W_2= W_0⊗{σ2}，W_3= W_2⊗{σ3}。由于要满足 4 个动作执行的先决条件，最后得到的 $\mathcal{M}4$=<{w_{D5}}, {(D5, D5), {(D5, D5)}, $[\ ]_{\mathcal{M}4}$>；而$[D5]_{\mathcal{M}4}=[D5]_M=\{w_{D5}\}$（更新并不改变原来状态中的事实，即原子句子的赋值），所以我们有$\models_{\mathcal{M}4}K_aD5$ 和$\models_{\mathcal{M}4}K_bD5$，而且只有 D5 满足条件，它也就是所要求的结果。 ┤

4 结论与展望

我们对一类涉及知识的推理难题构造了认知模型，并且用公共宣称的转换模式和更新模型分别解释了其中的推理结构，对更新过程进行了细致的形式描述。通过本文可以看出，“转换模式”和认知更新模型对这些日常生活中某些涉及知识变化的直观推理难题进行的刻画是非常精确的。对生活中某个局部知识更新的直观意义的描述也是非常清晰的。在本文中，那些涉及“不成功更新”的推理难题得到了严格地解决，并且某些隐含的问题被挖掘出来，比如冗余性问题和时态问题等。在上述动态模型下，这类推理难题的结构被清晰地展现出来。

然而，值得注意的是，在 3.3 的 $\mathcal{M}2$ 中，我们规定它的状态域为 W_0 而非 W_1。这主要是

因为σ2 涉及过去时态。直观上这样的解释说得过去，但我们并没有在 pre(σ2)体现出来，形式上并不完美。对于含有时态宣称的形式化应该是我们下一步的任务。由此带来的结果或许会对现有的动态认知逻辑和更新逻辑的内容产生一定的影响。

感谢

本文在写作过程中，鞠实儿, van Benthem，李小五，刘奋荣，欧佳致同作者进行了讨论，提出了有益的建议，在此一并感谢！

参考文献

[1][AGM85]: C. E. Alchourrón, P. Gärdenfors and D. Makinson, On the logic of theory change: Partial meet contraction and revision functions, *J. Symbolic Logic* **50** (1985), 510–530.

[2][BMS03]: A. Baltag, L. Moss & S. Solecki, 'The Logic of Public Announcements, Common Knowledge and Private Suspicions', Update version of [BMS98], 2003.

[3][BMS98]: A. Baltag, L. Moss & S. Solecki, 'The Logic of Public Announcements, Common Knowledge and Private Suspicions', *Proceedings TARK 1998*.

[4][BPX01]: N. Belnap, M. Perloff & M. Xu, *Facing the Future*, Oxford University Press, Oxford 2001.

[5][vBen05]: J. van Benthem. Open Problems in Logical Dynamics, Report PP-2005-06, ILLC Amsterdam 2005.

[6][vBL04]：J. van Benthem & F. Liu, 'Diversity of Logical Agents in Games', Report PP-2004-13, ILLC Amsterdam 2004.

[7][vBen03]: J. van Benthem,, 'Rational Dynamics and Epistemic Logic in Games', in S. Vannucci, ed., *Logic, Game Theory and Social Choice III*, University of Siena, department of political economy, 19–23, 2003.

[8][vBen01]: J. van Benthem. 'Games in Dynamic Epistemic Logic', *Bulletin of Economic Research* 53:4, 219–248, 2001.

[9][BdRV01]: P. Blackburn, M. de Rijke & Y. Venema, *Modal Logic*, Cambridge University Press, Cambridge 2001.

[10][vDK04]: H. van Ditmarsch and Barteld Kooi. Unsuccessful Updates, in Section A.2 *Philosophical Logic (a course introduction)*, University of Otago 2004.

[11][vDit00]: H. van Ditmarsch, Knowledge Games, dissertation DS-2000-06, Institute for Logic, Language and Computation, University of Amsterdam & Department of Informatics, University of Groningen, 2000.

[12][FHMV95]: R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about knowledge*. MIT Press, 1995.

[13][KHT00]: D. Kozen, D. Harel & J. Tiuryn, *Dynamic Logic*, MIT Press, Cambridge 2000.

[14][KHT84]: D. Harel, D. Kozen & J. Tiuryn, Dynamic Logic, in *Handbook of Philosophical Logic Volume {II} --- Extensions of Classical Logic*, by D. Gabbay and F. Guenther, eds., 497-604, D.~Reidel Publishing Company: Dordrecht, The Netherlands 1984.

[15][McC78, 81]: J. McCathy, Formalization of Two Puzzles Involving Knowledge, Stanford University, http://www-formal.stanford.edu/jmc/, 1978-1981.

Formalization for a Kind of Logical Puzzles Under the Framework of Update Models

Jiahong Guo

(Philosophy Department, Sun Yat-sen University, email: jiahong@pitt.edu)

Abstract: In the last three decades, many logicians have become interested in actions played by agents and processes of cognitive reasoning. Dynamic structure is now one of the central themes of contemporary logic. There are lots of subjects such as dynamic epistemic logic, update logic, belief revision, branching time and etc. which represent the dynamic trend. This article tries to use dynamic structure from “the logic of public announcement” (one kind of dynamic epistemic logics) and update product models to finely explain some type of logical reasoning puzzles. Those puzzles contain knowledge update and actually unsuccessful updates are concerned. Under the formalization of the above models, those puzzles can be resolved easily and strictly, and some problems such as redundancy and tense are clarified. Meanwhile, this also shows that the models we apply do capture some parts of the concept of knowledge update we use in our daily life.

Key Words: Public Announcement, Reasoning Puzzles, Unsuccessful Updates, Update Models