

用游戏解释语言：博弈语义简介

罗生门，清华大学哲学系本科生，主要兴趣为自然语言语义学，同时对语言哲学和认知科学感兴趣。

语言的意义[1]是如何生成的？在语言学家和逻辑学家的传统观点中，基本语言片段的意义由词库指定，在一门语言中相对固定，幼儿通过学习这个词库来掌握意义，而复杂片段的意义则由其组成部分的意义决定。这种观点下，意义不涉及主体的行为，可以称为“零主体的意义”。而从上世纪七十年代开始，将意义视为信息交互的观点开始兴起：言说者通过语言改变听者头脑中的信息状态，而意义便由这种变化确定：一句话的意义就是其能给接受该句子的人带来的信息变化。这种观点下，意义涉及言说者的目的与行为，因此可以称为“单主体的意义”。兴盛于阿姆斯特丹的动态语义学（dynamic semantics）即以此为基础建立，传统的语义学-语用学的分野在这个框架下重新划定，自然语言在其中也被看作是用以改变主体状态的编程语言。

一旦将信息交互纳入考虑范围，我们自然可以更进一步：既然只有当一句话对听者引起了言说者所希望的信息改变时，交流才是成功的，而听者对这句话又可能有不同于言说者所希望的解读，因此可以想象，在语言互动的过程中，言说者在措词造句时，需要考虑听者会如何解读（“我需要知道你知道什么”），而听者在解读时也会相应地考虑到言说者会如何考虑自己如何解读（“我需要知道你知道我知道什么”），以此类推。语言的意义，便隐伏在如此织就的信念、知识与信息之网中，这便是“双主体的意义”[2]。

可能这样的引入有些抽象，下面的两个例子或许能展现一些知识与语言意义交互的风味。考虑以下情景：某次战役中，两个将领A和B分驻敌军阵地两侧，两人都知道各自的部队单独进攻都会被歼灭，只有同时进攻才能取胜。由于无线电通讯被干扰，两人事先约定凌晨四点一起进攻。即将发动进攻时，A发现情况有变，于是派人往B驻地报信：计划有变，听到任何人的枪响后一起进攻（这条消息简记为 atk ）。B得知消息后心想：现在A并不能确定我是否知道 atk ，因此也不敢发起进攻。于是B也派人向A回信。A得知消息后心想：现在B并不能确定我是否知道B知道了 atk ，因此即使听到枪响也不敢发起进攻（枪响可能是敌人发出的，因此不能认为B能如下推理：“如果A不确定，A就不会开枪，因此只要有枪响，就说明A已经收到了回信”）。于是A又派人向B回信。B得知消息后心想：现在A并不能确定我是否知道了A知道了我知道了 atk ，因此……如此下去，两人永远无法开始进攻（如果出于意外，有人开了枪，那么A、B两方都会因为无法确定对方会进攻而逃跑，这种无法确定是出于对对方认知状态的了解（或不了解），而非不够信任对方）。在这里，由于信息交互的不完整，A的枪响无法传递任何意义，即使所有人都能听到。

考虑另一个情景[3]：在一个英国小镇上，Johann是唯一的德国人。一天他去喝啤酒，刚进酒吧门口，Martin便大声笑道：“Hans又来喝啤酒了！”众人皆笑。镇上没有人叫Hans。但这句话成功地传递了应有的意义，每个人都知Hans这个名字说的是Johann。这是如何做到的？正是由于Martin知道镇上没有人叫Hans，且每个人都知道Martin知道镇上没有人叫Hans，且Martin知道每个人都知道Martin知道镇上没有人叫Hans，且……因此，所有人都知道Martin必然另有所指，而最合适的对象便是Johann（为什么Johann最合适？因为Martin知道Hans是常见的德国男名而Johann是德国男人且常来喝啤酒且刚刚进酒吧，且每个人都知道Martin知道这些，且……）

可以看出，意义和语言群体中的信息交互、合作与公共知识有着密切的关系，而这些主题能够在博弈论中自然地表示。因此，我们有了一种别于传统语义学的思路：用博弈确定意义。这看似新奇，事实上在晚期维特根斯坦的语言游戏中便有朴素的呈现，甚至Peirce就已经提出过可以把量词解释为双人博弈了。我们可以用不同方式将这一思路发展成型，在其中一种路线中，我们把每个句子都看作一场博弈，用虚拟玩家的必胜策略存在与否确定句子的真

值。本文即是对这种博弈语义 (game-theoretical semantics) 应用于自然语言的初步介绍, 其中第一节介绍用博弈/游戏解释语言意义的哲学背景, 第二节说明用博弈语义解释可能优于经典语义的一些自然语言现象, 第三节介绍博弈语义如何应用于自然语言中的量词和回指现象, 第四节介绍其他用博弈解释语言意义的方式。

语言游戏: 另一种语义

考虑下面两个问题: 语言片段和其对应的现实世界是如何连接的? 我们语言使用者又是如何确定语言表述的含义? 早期维特根斯坦用图像论回答了前者, Tarski语义 (一阶逻辑的经典语义) 则回答了后者。Stenius (1960) 将图像论表述为“世界的成分被与其逻辑类型相同的语言实体表示” (the ingredients of the world are represented in the language by entities of the same logical type)。注意这里的“语言实体”不是“语言符号”, 而是“个体”、“ n 元关系”这样的对象。在此我们需要区分三个不同的概念[4]: 世界中的关系 (所有满足条件的 n -元组的集合)、语言中的关系、语言中用来表示关系的符号。前二者的区别在于一个以世界中存在的实体为论元, 一个以语言片段为论元, 而它们都可以被语言中的符号表示。如假设 R 是逻辑语言中的一个二元关系符号, Rab 成立当且仅当 a 代表的实体和 b 代表的实体满足 R 代表的关系, 那么 R 便定义了一个世界中的二元关系 R_1 。但同时 R 也定义了一个语言中的关系 R_2 , 它的论元是“ a ”“ b ”这样的逻辑语言中的变元。维特根斯坦的图像论说的是, R_2 (语言中的关系) 表示 R_1 (现实中的关系), 并且二者逻辑类型相同。而Tarski语义说的则是, R (语言符号) 表示 R_1 (现实中的关系)。

Hintikka (1977) 指出, 图像论和Tarski语义非常相似。首先, 根据两个理论得出的语言中的句子为真的条件完全相同, 这可以从上面的例子中看出。更重要的是, 两种理论都把语言和其对应的世界之间的表示关系当作一种基本概念, 不再进一步分析 (用第一段的说法, 它们都是“零主体的意义”)。它在《逻辑哲学论》中被称为“命名”, 在Tarski语义中则被称为“解释”。然而, 这种表示关系是需要分析和解释的: 从引言中的两个例子就能看出来, 学习语言和符号的人不可能单单观察语词和世界就能总结出其间的表示关系。他们必须观察语言使用者怎么使用这些词——观察由特定规则约束的语言运用。在晚期维特根斯坦看来, 正是这种运用建立并维持了语言和世界之间的表示关系。一根带有划痕的棍子和测量尺有什么区别? 后者之所以是后者, 就是因为它被以特定方式运用。同样, 符号或声音的组合也因为被运用而成为语言, 这种运用便被称为语言游戏。

Stenius (1967) 根据自己的理解设想了一个简单的语言游戏: 假设有个园丁 A 和她的助手 B , 她们负责的花床分成许多小块, 每一块上都种了花。首先, 每天 A 和 B 都去检查花床, 然后 B 在纸上记录若干符号给 A 检查: Pa, Pb, Qc, Pd , 等等。这些符号被用来表示开花状况: a, b, c 表示小块的名字, P, Q 表示开花与否。但这些预期中的意义并不是由 A 直接告知 B 的, A 只根据花床的实际情况告诉 B 每个记录是对还是错。当 B 通过这种方式学会了符号的运作后, A 就不再需要每天去花园了, 只需要阅读 B 每天的纸条就能得知花园里的信息。这便是可以产生意义的语言游戏。

语言游戏不是语言中 (关于目标语言) 的游戏, 而是通过语言进行的游戏, 往往涉及非语言的语境, 比如上文中的开花情况。而语言片段的意义由其在这种游戏中的角色确定。语言学习者要确定一个语言片段的意义, 只需要观察游戏玩家对其如何反应。如何验证“房间里有两个球”这句话? 很简单, 到房间里去, 看看能不能找到两个球。可以看到, 这种语义学观点引入了行动者, 它们通过特定规则的行动来验证句子是否成立。沿承这一思路, 我们便可以发展出博弈语义。

自然语言的信息独立

本节将介绍自然语言中的信息独立现象, 由此说明**IFL** (IF-逻辑) 比**FOL** (一阶逻辑) 更适合刻画自然语言。由于博弈语义能比Tarski语义更好地解释**IFL**, 因此博弈语义在这些现象上也是自然语言语义的更好选择。

在当代语言学 (生成语言学) 中, 我们通常认为每个自然语言句子在大脑处理过程中都会产生一个语音形式

(phonetic form) 和一个逻辑形式 (logical form)。语音形式就是我们在语言交流时会实际说出来的句子，而逻辑形式则是用来解读意义的形式，这两种形式之间通常并不一致。比如，“I love every student”是一句话的语音形式，其逻辑形式可能是“Every student (1) [I love t_1]”。在本节中我们将假设每个自然语言句子都可以通过其逻辑形式和一个逻辑语言句子对应，并讨论后者的博弈语义，从而绕过对自然语言句子的直接语义分析。

否定移位

考虑以下句子：

Abdullah does not believe that Hiromi is at home.

母语者通常会将其解读为“Abdullah相信Hiromi不在家”，而非“Abdullah不相信Hiromi在家”。这两种解读并不等价，比如Abdullah可以对Hiromi在不在家保持中立：既不相信她在家，又不相信她不在家。两种解读可以分别用如下两个引入信念算子的FOL句子刻画：

$$(a) B_{Abdullah} \neg Hh$$
$$(b) \neg B_{Abdullah} Hh$$

其中， Hh 表示Hiromi is at home。由于英语原句的语音形式中“not”在“believe”之前，而其逻辑形式是(a)，因此这里似乎发生了否定移位：否定词本来在模态算子辖域内，而在语音形式中移出了；或者说其本来在模态之外，而在逻辑形式中移入了。但这个现象同样可以用信息独立解释：假设这里的模态算子受某种影响独立于否定词，即 $\neg(B_{Abdullah} / \neg)Hh$ 。考虑这个句子代表的博弈：首先双方互换角色，然后在无视这一互换的情况下，证伪者选择一个Abdullah认为最可能的情况，如果此时Hiromi在家则Abdullah获胜。在这个博弈中证实者什么都不用做，因此其有必胜策略当且仅当在所有Abdullah认为最可能的情况中Hiromi都不在家，显然这和(a)是真值等价 (truth equivalent) 的（它们并非假值等价 [falsity equivalent]，但这通常不在我们的考虑范围内），因此信息独立可以不用移位解释原句中的现象。

如果两种说法都能解释这一例子，为什么要把这看作自然语言中存在信息独立的证据呢？让我们考察下面的句子：

Nobody doesn't believe that Hiromi is at home.

显然，这句话最自然的解读等价于 $\neg(\exists x)\neg B_x Hh$ 。然而，否定移位假说所预测的逻辑形式为 $\neg(\exists x)B_x \neg Hh$ ，这和自然解读并不等价，因此否定移位假说无法解释此处否定词为何不移位了。而信息独立假说则可以预测这一现象，只需要认为模态算子独立于两个否定词： $\neg_1(\exists x)\neg_2(B_x / \neg_1 \neg_2)Hh$ 。注意到该句真值等价于 $(\forall x)B_x Hh$ ，因此，我们认为信息独立更能解释英语中否定词的表现。

模态不透明：De re - de dicto之别

自从弗雷格开始，语言哲学家们就关注到了一个问题：通常来说，如果两个词组指代的是同一个东西，那它们在语言片段中是可以互相替换而不改变句子真值条件的。比如，“我有三本书”、“本文作者有三本书”、“我有 (2+1) 本书”、“我有三本特定类型的纸质出版物”这四句话是等价的，只要其中被替换的词语的指称相同。然而这在被模态限定的句子中不一定成立。比如，我们通常认为在现实世界中，“崇祯帝知道了北京陷落之事”是真的，但“崇祯帝知道了北平陷落之事”则不一定。尽管北京就是北平，但由于崇祯帝并不知道“北平”这个词的所指，因此这两个词不一定能相互替换。可以说，模态算子创造了一个对替换“不透明”的环境，这为歧义创造了空间。考虑下面这个句子：

小李知道老牛在寝室。

这句话有两种可能的解读。假设老牛是小陈的班主任，而小李不知道此事。设想两种情景：(a)小李本来不认识老牛，回寝时在门口看到了一张掉落的名片，上面有老牛的名字，又听到门内传来说话声，因此知道老牛在寝室；(b)小李和老牛在寝室里面对面聊天，因此知道老牛在寝室。两种情况中“小李知道老牛在寝室”都成立，但在(a)中“老牛”和“小陈的班主任”不能替换，因为小李知道的只是“老牛”这个名字而不知道具体的人；(b)中则可以互相替换，因为小李知道的是某个特定的人在寝室，至于这个人用什么方式指称、小李是否知道这些指称并不重要，小李甚至不需要知道和自己说话的人是老牛也能使得“小李知道老牛在寝室”在这种解读下为真。

我们把(a)这样的解读称为*de dicto*解读，(b)称为*de re*解读。前者字面义是“关于说法”，而后者字面义是“关于对象”。我们可以用两个不同的逻辑句子翻译这两种解读：

$$(a) K_{Li} Qn$$

$$(b) (\exists x)(x = n \wedge K_{Li} Qx)$$

其中， n 表示老牛， Qx 表示 x 在寝室。(a)的翻译并没有什么问题，但(b)的翻译如果要用来解释自然语言的语义，则有一个重要缺陷：“小李知道老牛在寝室”这句话中找不到对应存在量词的句法证据。而博弈语义和信息独立则可以轻易地解释这个现象：

$$(b') K_{Li} Q(n/K_{Li})$$

这样，(b')说的就是：不论在何种小李可以设想的情况下（不管小李对面这个人叫不叫老牛、是不是小陈的班主任），都能选出一个固定不变的个体（也就是小李对面这个人）填入“小李知道老牛在寝室”中的“老牛”处。相应地，(a)说的则是：不论在何种小李可以设想的情况下，都可以根据这个情况选出一个可变的个体（这张名片的主人）填入“老牛”处。显然，这两个句子分别对应两种解读（读者可能还会注意到，(a)和(b')是假值等价的，也就是说两种解读为假的情况相同）。Rebuschi & Tulenheimo (2011)指出，如果把“知道”换成“相信”这样非事实的模式算子，还可能出现第三种解读，介于*de dicto*和*de re*之间，可称为*de objecto*。

现实算子与可能世界回退

我们把“其实”、“现在”这样的词称为现实算子（actuality operator）。考虑以下两个句子：

老牛相信有人要害他，但那些人其实是在帮他。
我曾以为我现在不会学哲学。

请读者尝试用带模态算子的FOL刻画这两个句子。读者会发现有个问题：以第二句为例，我们需要首先说存在某个过去的情况，彼时如何如何。但这个“如何如何”说的却是现在的状况——如果仅仅是“过去的未来”，那自然可以简单地用 $PF\phi$ 描述，但现在我们需要从过去精确地跳回到现在。同样，第一句也同样需要从老牛认为最可能的情况精确地跳回到现实状态，而我们通常使用的模态算子无法做到这点。这就是现实算子导致的问题。

对此，语言学家和逻辑学家们提出了一些解决方案。其中一种是认为存在能回退到之前情况的回退算子（backwards-looking operator）。但这类算子本身在自然语言中使用很局限，并且缺乏句法证据。同时，在一些复杂的句子中回退算子的语义会显得十分臃肿，如“老牛相信小李觉得有人要害他，但那些人其实是在帮他”这句话中需要回退两次。另一种解决方案是改变我们的元逻辑设定，认为在衡量一个句子在一个情况下的真值时，我们还需要考虑其在我们“到过”的情况中的真值，这被称为模态的多维语义（multi-dimensional semantics for modal concepts）。但博弈语义和信息独立可以提供更简洁的解释：我们只需要认为，在衡量“我不会学哲学”时证伪者不知道此前证实者在 P 处选择的情况即可。

在对解释同一现象的诸多理论进行讨论时，我们往往会认为博弈语义和信息独立之外的理论缺乏句法上的证据。那信息独立的句法证据呢？很可惜，目前也没有！然而，博弈语义不是组合的（compositional），因此并不严格要

求其与句法结构需一一对应——当然，没有句法证据总归是一大缺憾。不过，至少信息独立能够将几种看似无关的现象用同样的方式解释，其他没有独立证据的假说相比之下则更为*ad hoc*。

自然语言的博弈语义

本节主要讨论量词和回指代词的博弈语义，它们不论在哪种语义中都是最重要的词类。Hintikka对英语中量词的博弈规则规定如下：以“some”为例，由于英语中不存在 $\exists xYx$ 这样的结构，我们将“some X that Y ”视作一个整体代入单个个体名，并用另一个分句来处理其内部结构，即：

当博弈进行到 $G(A - \text{some } X \text{ that } Y - B; \mathcal{M})$ 时，这一回合由证实者从 \mathcal{M} 的论域 M 中选择一个对象代入“some X that Y ”。博弈继续，假设所选的对象在目标语言中名为 b ，则下一回合的博弈局面为 $G(A \ b \ B, b \text{ is an } X, \text{ and } b \ Y; \mathcal{M})$ 。把所选的对象加入选择集。

何谓“选择集”下文会谈到。其他量词如“any”，“every”，“all”，“a(n)”等规则类似[5]，区别主要在行动者是证实者还是证伪者，证实者的回合后用“and”连接新增分句，证伪者则用“if”，正如FOL中表示“所有 P 都是 Q ”的句子是 $\forall x(Px \rightarrow Qx)$ 而非 $\forall x(Px \wedge Qx)$ 。以下面这个句子为例：“Some student who hates every professor respected a lecturer”运用三次量词博弈规则后可以得到“Gabriel respected Martha, Martha is a lecturer, Gabriel is a student, and Gabriel hates Evgeny if Evgeny is a professor”。至于“and”等命题连词，也可以用与逻辑语言中类似的博弈规则处理。

容易看出，以何种顺序处理上句中的三个量词并不紧要。因此，同一个自然语言句子可能以不同的顺序运用博弈规则处理各个成分，因此有时需要额外的规则（比如必须从左到右）来限制解读顺序。这和逻辑语言不同：一个FOL或是IFL句子只有一种解读顺序（即唯一可读性），顺序是由括号确定的。而自然语言中没有这样的括号，量词与连结词虽然有约束辖域（其内的成分受该量词或连结词约束），但没有顺序辖域。

以上部分基本只是将FOL的博弈语义照搬到自然语言的对应部分中，接下来我们来考察一些自然语言独特的组分。首先考察代词。传统语法认为代词有两种用法，如果代词所指代的事物由上下文确定，则称为回指（anaphora），如果由语境而非上下文确定则称为直指（deixis）。例如，“老牛把他的钱都花光了”这句话中的“他”可以指老牛也可以指别人，指老牛则为回指，指别人则为直指。

形式语义学通常把回指代词处理为约束变量（这一说法并不精确，回指代词还有其他情况，我们在此不作细分），这是因为形式语义学的基本原则是组合性，约束变量是能满足组合性且解释回指代词用法的最佳选项。然而，既然我们的博弈语义是非组合的，自然也可以用其他方式解释回指代词。大体来说，我们将其视作确定描述（definite description）：“he”看成“the male”，“she”看成“the female”等等，并大致采用罗素对确定描述的分析：“The cat meows”等价于“There is a cat and there is only one cat, and it meows”。不同的是，代词并不在模型的整个论域中选取个体（如直指代词那样），而是只考虑选择集（choice set）：在博弈中之前的回合里被选取过的对象构成的集合。于是，以“she”为例，我们有如下博弈规则：

当博弈进行到 $G(A - \text{she} - B; \mathcal{M})$ 时，这一回合由证实者从此时的选择集中选择一个对象代入“she”，同时证伪者也从其中选择一个对象。博弈继续，假设证实者所选的对象在目标语言中名为 c ，证伪者所选的对象名为 b ，则下一回合的博弈局面为 $G(A \ c \ B, c \text{ is a female, and if } b \text{ is a male then } b \text{ is the same as } c; \mathcal{M})$ 。

如果把语篇中的每个句子看成语篇的子博弈，那语篇回指（即不同句子之间的回指）也能以相同的方式解释。这种处理方式有一些优点，比如在解释“A couple were sitting on a bench. Suddenly he got up.”时，用博弈语义可以轻松解释为什么这里的“he”能被自然地理解，还能据此反推出这对“couple”应当是一男一女。又如，“Magdalena went to her office. Aliya did too.”这句话既可以指两人分别去了自己的办公室，也可以指两人都去了Magdalena的办公室。解释这类现象通常需要认为回指代词具有两种不同用法，而博弈语义则不需要这一预设也能轻松解释：博

弈进行到前一个“her”时，选择集还只有代入“Magdalena”的一个个体，因此前一个“her”只能指代Magdalena；进行到后一个“her”时选择集已经有了两个个体，因此有两种可能情况（当然，也有可能两个“her”指代的都是第三人，但这是代词的直指用法，不在此处的讨论范畴）。

一旦我们把回指看成是从选择集中选取个体代入，那么我们就拓宽“回指”这个概念的范围。比如，限定描述也可以分成直指（“*The present king of France is bald*”）和回指（“*She kept a cat and a dog, and I was afraid of the dog*”）两类，后者可以与回指代词被同样解释：只需将上述关于“he”的博弈规则中的“he”替换为“the X”，把新增的分句中“*b is a male*”替换为“*b is an X*”即可。甚至解释回指量词也完全可能，比如下面这个语篇：

三个学生参加了考试。有一个顺利通过了。她显然做好了准备。另外两个则没有及格。

可以很清楚地看到量词的语篇影响和语篇回指是如何自然地实现的。“三个学生”是由证实者选择对象代入并放入选择集中，“有一个”、“她”是从选择集中选取，“另外两个”则取决前两个量词的选择，因此依赖于前两个量词。然而，这样的解释有一个明显的漏洞：如果把上面的“三个”换成“所有”，那么选择集中应该只有一个对象，那就是由证伪者选择并代入“所有学生”的个体。然而，这样的话“有一个”的解释将会偏离我们希望的自然语言解读，而“另外两个”则根本无法解读，因为选择集中只有一个个体。因此，我们可以对规则打个补丁：区分单句选择集和语篇选择集。前者只对本句有效，而后者对整个语篇有效，语篇回指词只能从后者中选择个体代入。这样，之前提到的所有博弈规则中，选择的对象都只放入单句选择集。如果这个对象是由证实者选择的，那么他同时被传递到语篇选择集中；如果是由证伪者选择的，那么可供证伪者选择的所有个体都被传递到语篇选择集中。这也解释了如下几个（与donkey sentence同类的）语言事实（其中句子前的*表示该句不合语法）：

*Every student thinks she's smart. She is self-confident.

Every student thinks she's smart. They are self-confident.

*Mary didn't see some student. He was hiding in the corridor.

Mary didn't see some student. They were hiding in the corridor.

更多视角，更多博弈

以上便是逻辑语言和自然语言的博弈语义的基本内容。我们在本文的开头提到，博弈语义只是用博弈解释语言现象的一种方式，那么在这最后一节，就让我们看看另一种途径：由David Lewis提出的信号博弈（signaling game）。

考虑以下情景：一个卡车司机A正在倒车，其助手B在车后指挥。可能由于环境过于嘈杂，他们选择采用手势交流：B根据是否还能继续倒车打出不同的手势，A根据看到的手势选择是否继续倒车。由于两人都可以选择对于特定的输入作何种输出，这个场景很容易被形式化成一场博弈：对于情况集合 $T = \{\text{可以倒车, 不能倒车}\}$ ，信号集合 $F = \{\text{手势1, 手势2}\}$ 和行动集合 $Act = \{\text{倒车, 不倒车}\}$ ，B可以选择策略函数 $f_B : T \rightarrow F$ ，A也可以选择策略函数 $f_A : F \rightarrow Act$ 。A和B各有4种策略可以选择，总共有16种可能情况。

由于这个博弈中双方的目的相同，因此可以认为如果事实上能倒车并且A倒了车，或者不能倒车且A没有倒车，那么A和B都得1分，除此之外的局面不得分。这样我们可以计算出选择每种策略组合时每种情况下的两人分数，进而通过两种情况的分布概率算出每种策略组合的期望得分。容易想到，只有两种策略组合能满足纳什均衡，一种是 $f_B = \{(\text{能倒车} \rightarrow \text{手势1}), (\text{不能倒车} \rightarrow \text{手势2})\}$ ， $f_A = \{(\text{手势1} \rightarrow \text{倒车}), (\text{手势2} \rightarrow \text{不倒车})\}$ ，另一种是将手势1与手势2对调。在这两种策略下，双方的得分期望都应该是1，因此选择哪种都有可能。如果我们将打手势的成本考虑进来，假定手势1比手势2更简单，并且假定能倒车的概率小于不能倒车，那么尽管仍然有两个均衡策略组合，但用手

势1对应不能倒车、手势2对应能倒车的组合中双方的得分期望会更高，为帕累托最优，我们认为司机和助手会采用这种组合。

博弈可以产生意义——这在信号博弈中再次得到了呈现。如果我们比较信号博弈和博弈语义，会发现它们所涉及的博弈-语言意义之间关系并不一样。首先，两种博弈的性质不同：在信号博弈中，博弈是实际发生的，博弈玩家就是语言使用者。而在博弈语义中，博弈玩家是虚设的，博弈也未曾实际发生，语言使用者所做的不过是声称在某个博弈中某个玩家存在必胜策略。因此，维特根斯坦所说的语言游戏更接近信号博弈（司机-助手的例子和前文中园丁-助手的例子也很相似，只是后者是园丁通过语言游戏让助手学习意义，而前者是双方共同“约定”意义）。

其次，它们处理的是两类不同的语言问题。博弈语义中事实上并没有直接的博弈-语言交互，博弈分析的是句子结构的逻辑骨架，逻辑才是分析语言本身的方式。相应地，这种博弈只有在对应的逻辑操作（如命题连词）的意义（也就是博弈规则）和基本的词汇意义（如常量和关系谓词）都已经确定时才可能实现。而信号博弈则是被用来建立这两者的（至少是后者），此时博弈可以直接产生和分析语言本身，博弈-语言的关联也更为深刻。一些特殊的情况下（如语言不通），我们可能会需要信号博弈来产生意义，但在标准的交流情景中，可以认为上述有关意义确定的假定是成立的，整体（或至少是局部的）语言惯例已经形成，此时我们可以专注于更高阶的语言现象，也就是句子的真与假。

事实上，还有很多其他用博弈解释语言现象的方式。例如，在实际的语言交流中很多时候我们并不知道实际情况（模型）是怎样的，而是假设对方所说的都符合实际情况，据此构造一个模型，这个过程可以用模型构造博弈（logic games of model construction）解释。相应地，和模型论语义相对的证明论语义也提出了证明博弈（dialogue games for proof）。而在这些词项意义、句子真值、模型和推论都被确立了之后，真正的语言交流才刚刚开始。在开头我们提到过“单主体的意义”：把语义看作是对听者的信息状态的改变。改变牵涉到信息更新和信念修正，这一过程也可以用博弈解释。如果我们把视角放得更大，就会注意到自然语言的惯例事实上是在整个语言社群的基础上形成的，这种惯例形成后需要被重复无数次（无穷博弈），经受各种可能的变化而依然保持相对稳定。我们甚至可以设想，怎样的语言策略组合是进化稳定对策（evolutionarily stable strategy）？也许最微观的信号博弈机制就已经被蕴含在了这样的进化稳定对策里？如此种种，展示了逻辑-语言-博弈的关联可以在极多样的层面上体现，请读者以最开放的态度探索这些问题吧。

脚注

[1] 本文中不会严格区分意义和指称。

[2] 以上两段内容主要概括自van Benthem (2008)。

[3] 该例子来源于唐浩老师两年前在清华大学讲授《近代西方哲学》的课程讲义。

[4] 这个区分来自 Hintikka (1977)。

[5] Hintikka (1973)认为英语中有两种a(n)，一种是量词（“A student is here”），另一种是谓词的一部分（“Catherine is a student”）。后者被当成游戏终局，不需要进一步分析。见该文注21。

参考文献

Van Benthem, J. (2008). 'Games That Make Sense': Logic, Language, and Multi-Agent Interaction. in K. R. Apt and R. van Rooij (eds.). *New Perspectives on Games and Interaction*. Vol. 4. Amsterdam University Press, Amsterdam. 197-210.

Hintikka, J. (1974). Quantifiers vs. Quantification Theory. *Linguistic Inquiry* 5. 153-177.

Hintikka, J. (1977). Language-Games. *Dialectica*. 31(3/4), 225-245.

- Hintikka, J. and Kulas, J. (1983). *The Game of Language*. D. Reidel Publishing Company, Dordrecht.
- Hintikka, J. and Kulas, J. (1985). *Anaphora and Definite Descriptions. Two Applications of Game-Theoretical Semantics*. D. Reidel Publishing Company, Dordrecht.
- Hintikka, J. and Sandu, G. (1989) Informational Independence as a Semantical Phenomenon. in J. E. Fenstad *et al.* (eds.). *Logic, Methodology and Philosophy of Science*, Vol. VIII. Elsevier, Amsterdam. 571-589.
- Hintikka, J. and Sandu, G. (2011). Game-Theoretical Semantics. in J. van Benthem and A. ter Meulen (eds.). *Handbook of Logic and Language*. Elsevier, Amsterdam. 415-466.
- Majer, O., Pietarinen, A., and Tulenheimo, T. (2009). *Games: Unifying Logic, Language, and Philosophy*. Springer.
- Pietarinen, A. (2001). Most Even Budgeted Yet: Some Cases for Game-Theoretic Semantics in Natural Language. *Theoretical Linguistics*, 27 (2001), 20-54.
- Rebuschi, M. and Tulenheimo, T. Between *De Dicto* and *De Re: De Objecto* attitudes. *The Philosophical Quarterly (1950-)*. October 2011, Vol. 61, No. 245 (October 2011). 828-838.
- Sandu, G. (1993). On the Logic of Informational Independence and Its Applications. *Journal of Philosophical Logic*. Vol. 22, No. 1 (Feb, 1993). 29-60.
- Sandu, G. (2008). Games in Language. in K. R. Apt and R. van Rooij (eds.). *New Perspectives on Games and Interaction*. Vol. 4. Amsterdam University Press, Amsterdam. 179-196.
- Stenius, E. (1960). *Wittgenstein's Tractatus: A Critical Exposition of Its Main Lines of Thought*. Blackwell, Oxford.
- Stenius, E. (1967). Mood and Language-Game. *Syntese*, 17(1):254-274.