国家科学技术学术著作出版基金
NFAPST

陆汝钤 著

计算系统的形式语义 下

FORMAL SEMANTICS OF COMPUTING SYSTEMS

清华大学出版社

陆汝钤 著

计算系统的形式语义 下

FORMAL SEMANTICS OF COMPUTING SYSTEMS

清华大学出版社
北京

内 容 简 介

计算系统的形式语义是目前计算机科学理论研究的两大方向之一，其研究成果对程序设计语言、编译技术、应用软件、分布式系统等分支领域有重大的实际意义。本书大体上分为三个部分。第一部分是数学基础，为第一章。第二部分包括第二到第五章，概述了形式语义中的操作语义、指称语义、公理语义和代数语义四大经典流派。第三部分包括第六到第九章，概述了形式语义学的现代应用，分别介绍分布式系统、移动计算和移动通信系统、非规范进程代数和微观生命系统，以及量子程序设计语言的形式语义。

全书内容丰富，结构严谨，集形式语义学理论及其应用的有关分支之大成，系统地反映了这个领域各方面的研究成果，特别是它的近代发展潮流和趋势，并对不同流派的理论和方法给予了分析和评论。

本书可作为计算机科学专业研究生、本科生有关课程的教材或教学参考书，也可供有关专业或交叉学科的科研人员进修或作为工具书。

图书在版编目(CIP)数据

计算系统的形式语义/陆汝钤著. —北京：清华大学出版社，2017
ISBN 978-7-302-41494-0

Ⅰ.①计… Ⅱ.①陆… Ⅲ.①形式语义－研究 Ⅳ.①TP301.2

中国版本图书馆 CIP 数据核字(2015)第 212873 号

责任编辑：薛　慧
封面设计：何凤霞
责任校对：王淑云
责任印制：李红英

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，　http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　邮　　编：100084
社 总 机：010-62770175　　邮　　购：010-62786544
投稿与读者服务：010-62776969，　c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，　zhiliang@tup.tsinghua.edu.cn
印 装 者：北京雅昌艺术印刷有限公司
经　　销：全国新华书店
开　　本：170mm×230mm　　**印　　张**：118　　**插　　页**：10　　**字　　数**：2102 千字
版　　次：2017 年 1 月第 1 版　　**印　　次**：2017 年 1 月第1 次印刷
印　　数：1～1500
定　　价：398.00 元(全二册)

产品编号：012233-01

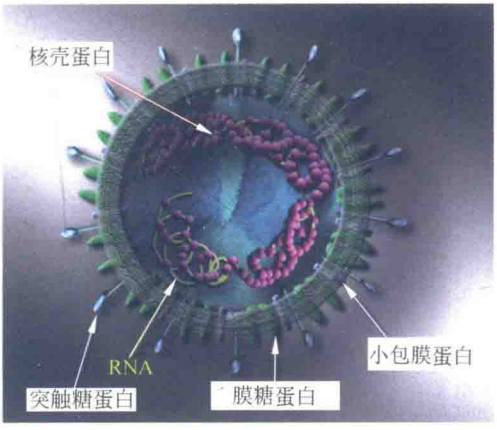

图 8.17.1　SARS 冠状病毒的结构

(译自 http://baike.baidu.com/picture/937450/937450/0/bd7faf355e4dafcba71e125c.html?fr=lemma&ct=single#aid=0&pic=bd7faf355e4dafcba71e125c)

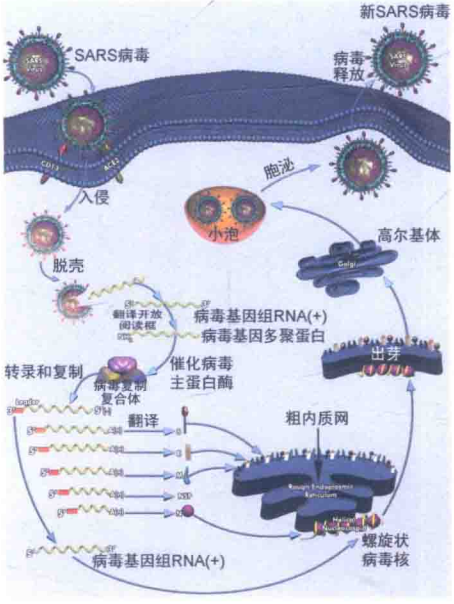

图 8.17.2　SARS 冠状病毒在细胞体内的自我复制过程

(译自 https://www.qiagen.com/us/products/genes%20and%20pathways/pathway%20details/?pwid=402)

马上相逢无纸笔，凭君传语报平安

——岑参

目　录

第 7 章　移动通信和移动计算系统的形式语义

7.1　概　　述

本章讨论进程代数在移动世界中的两个核心应用。一个是移动通信，另一个是移动计算。按通常的定义，移动通信是指一个通信双方可以在通信过程中移动其位置的无线网络，而无线网络是指不用物理连接（例如光缆）组织起来的通信网络（换句话说，是靠电磁波连接起来的）。最典型的例子就是手机通信。移动计算是指以下几种情况：一是正在进行中的计算在虚拟设备中移动。所谓虚拟设备，指的是同一个计算环境中通过程序分隔，不能随便交往的区域。防火墙就是一种典型的区域分隔，没有必要的口令和其他安全手段，任何程序不能随便穿过防火墙。二是正在进行中的计算在硬件设备之间移动。在分布式计算中早就有关于进程迁移的研究。现在网络发达了，这种计算的移动更是常事。尤其是在云计算环境中，计算在云中的移动成为云计算的一个重要特征。

为什么移动通信和移动计算会对进程代数的研究形成挑战（又是机遇）呢？我们首先来看一下移动通信的机制。持有手机的双方能够进行移动通信并非是因为无论通信双方之间的距离是如何遥远，它们能够始终保持直接联系。要知道手机的信号发射能力有限，无法发送到遥远的地方，使任意远处的任何人都可以收到。手机的通信实际上是要利用中介的，这些中介就是所谓的基站。通常每个基站管理的区域呈六角形蜂窝形式，所以也叫蜂窝基站，见图 7.1.1。这些蜂窝小区合在一起覆盖的区域就是移动通信有效的地域。每个手机用户一拨号，手机就开始搜索其所在的蜂窝区的基站的频率，一旦搜索成功，手机就和该基站建立联系，这个过程叫寻呼。联系建立以后，基站就把用户的需求（通话目

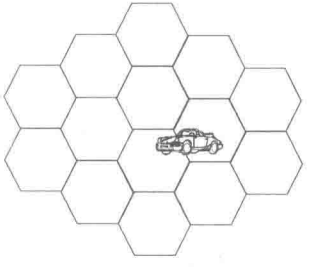

图 7.1.1　蜂窝小区和移动通信

标手机或座机)转发给基站控制中心,再由基站控制中心把信号转发给移动交换中心。根据用户的不同需求,如果目标也是手机则找到目标手机所在的蜂窝小区,再由目标蜂窝小区寻呼目标手机;如果目标是固定电话则转发给公共电话交换网,它会经过物理通信路线(如光缆)联系到目标座机。联系上后就可以开始通信,声音、文字或图像信号就这样源源不断地传往通信目的地。

当通信者离开当前小区进入新的小区时,手机和基站的连接会自动地从原来的基站转向新的基站。这个过程叫切换。切换的时机和切换到哪个基站一般根据信号的强度来决定。硬切换是先断掉与原来基站的联系,再建立与新基站的联系,这种连接一般会有短暂的中断,从而影响通信的质量。软切换先建立新基站的联接,再断掉原来基站的连接,新老基站的连接有一段共存的时间,通信质量比较高。

无疑,移动通信的特点会影响各类通信进程代数的设计。其中首先被考虑的是通信对象的切换。从上面描述的移动通信基站机制来看,有两个显著的特点:第一,通信双方至少有一方在(地理位置意义上的)移动中。第二,通信双方的任何一方可供选择的通信对象不是一个固定的集合,而是可以临时从外界获取新的通信对象地址,然后进行通信的。信道切换是后 CCS 时代进程代数新增的最重要功能。当然新一代进程代数还有其他很多技术进展,例如通信的私密问题,通信的质量(噪音)问题,等等。这些方面也有很多成果,但是与"移动"特征的联系已不是那么明显。更多的技术挑战,例如频道复用技术,则暂时没有在分布式通信进程代数的考虑之列。

传统的通信进程代数不具备解决上述问题的功能。首先,通信对象的切换问题就解决不了。在 CCS 中,通信港口都是固定的,写进程序里的。不能动态更换或创建新的港口。或者说,CCS 的通信网络是一个固定的拓扑结构。

意识到这个问题的人实际上很多,并且还先后提出了多种解决方案。其中最成功的是进程代数 π 演算,它是在 CCS 的基础上发展起来的,由 Milner 于 1990 年首先提出,在随后几年中逐步补充其内容,并于 1992 年由 Milner, Parrow 和 Walker 联袂发表详尽文章,才得以从此扬名天下。π 演算在其基本语法上非常像一个传值型的 CCS,因此它的功能也在许多方面和 CCS 很相像。为了节省篇幅,本节对这些相像的部分只作简略介绍,以利于突出 π 演算的特点。读者如果对一些基本的概念有问题,可以到上一章有关 CCS 的介绍中去寻找解释。

π 演算和 CCS 的重大区别在于:在 CCS 中数据名和港口名分属两个不同的集合,它们使用的标识符完全不同。而在 π 演算中不再有这样的区分,这两个集合合而为一,统称为名字集,它们统一代表信道(不再称港口)。所以在这个意义上,π 演算也可以称为传值型的名字演算。

π 演算和 CCS 的这个差别非常重要。正是因为数据名和信道名统一了，信道名就可以像(CCS 中的)数据名一样地被传送。于是，原本因没有配对的信道名而不能进行通信的动程之间，可以因为(在与别的动程的通信中)得到了信道名而从不能通信变成可以通信。π 演算描述移动通信的基本原理正来自于此。

直接从语法上看，π 演算的语法并不比传值 CCS 的语法更丰富。相反，在某些方面更加受限了。例如，传值 CCS 的布尔表达式在这里简化为只允许等于比较。相应的语法公式是：

$$[x = y].P \quad \text{和} \quad [x \neq y].P$$

它们的语义是(而且有的作者也这样写)

$$\text{If } x = y \text{ then } P \quad \text{和} \quad \text{If } x \neq y \text{ then } P$$

这个简化是很自然的，因为 π 演算中除了名字以外没有其他数据类型。那么把两个名字作大于或小于比较，或把一个名字和 0 作比较就没有意义了，更不可能把两个名字相加。唯一可能的比较就是相等或不等比较。另外，π 演算不需要 If $x=y$ then P else Q 这样形式的动程，因为此功能可以由 $[x=y].P+[x\neq y].Q$ 体现。

除了一些技术细节之外，π 演算有四个值得一提的特色或创新，它们对后来的进程代数产生了深远的影响。第一个特色就是把数据类型限定为只有“名字”这一种，从而统一了信道和信道传输的值，产生了“通信信道可以动态设定”的重要效果。因而我们也可以用“通信名字演算”来称呼 π 演算。对此，Milner 是颇有雄心的。他曾设想把名字计算与数值计算和符号计算并列，成为第三种主流计算理论。他还想把 π 演算做成一种分布式的 λ 演算。第二个特色是动程的屏蔽处理比原来复杂了。在 CCS 中屏蔽只针对信道名，而与信道传输的值无关。只要信道被屏蔽，传什么值都不行。这很简单。但是在 π 演算中，凡名字都有可能被屏蔽，无论它是作为信道名出现，还是作为被传输的名字出现。因此必须有相应的规则仔细地区别对待各种屏蔽情况。尤其是 π 演算增设了“约束输出”这一种传输类型，使传统进程代数中固定不变的作用域(由屏蔽操作设定)成为动态可变的，为研究分布式通信中的信息安全理论给出了一种理论工具。第三个特色是引进了“早语义”和“迟语义”的概念，与 CCS 中“强等价”和“弱等价”的概念互为犄角，大大丰富了进程代数语义的研究，并在实际上导致了后来“符号语义”的产生。第四个特色与上面三个特色有关，这就是引进了一大批丰富多样的双模拟语义，形成了一种绚丽多彩的双模拟理论体系。但梁园虽好，终非久恋之家。双模拟理论并不是研究进程代数语义的最终目的。人们真正感兴趣的是研究在什么条件下两个动程的行为不可区分，这就是等效的概念。它比双模拟要求更严。等效的一定双模拟，而双模拟不一定等效。为了使这个相对抽象的概

念具体化，可测试化，人们又引进了观察等价、测试等价等一系列手段。而双模拟理论虽丰富，却主要是为论证这些等效性和等价性而设计的技术措施。下面分别说明 π 演算的这些特色。

对于第一个特色，可以回顾一下 CCS。在区分信道和值的 CCS 中，信道名在动程运行过程中不能被更改，也不能被传递，只有值可以被计算，可以被传递，但是值不能当作信道用。在 π 演算中信道和值被统一成名字，名字除了不能计算外，具有 CCS 中信道和值的所有功能，一身而兼二任。因此很好理解为什么信道名可以动态地任意更换：只需在通信过程中把想要作为信道用的名字输送到信道的语法位置上去就行了。下面，我们就以移动通信功能为目标，看一个模拟手机和基站通信的例子，考查一下信道名动态设置的好处。

例 7.1.1　令基站动程为：

$$P_1 \overset{\text{def}}{=} \bar{a}b.\,b(c).\,b(u).\,\bar{c}u.\,P_1 \quad \cdots \quad P_n \overset{\text{def}}{=} \bar{a}d.\,d(e).\,d(t).\,\bar{e}t.\,P_n$$

手机发信动程为：$R \overset{\text{def}}{=} x(w).\,a(z).\,\bar{z}w.\,m(n).\,\bar{z}n.\,R$

手机通过基站发送信息的移动通信动程为：$P \overset{\text{def}}{=} P_1 \mid \cdots \mid P_n \mid R$

当手机 R 需要发送信息时，手机使用者（看作一个未明显给出的动程）先通过信道 x 向 R 输入对方地址 f，“存于”w 中（此时名字 w 变成名字 f）。然后手机接收基站信息。共有 n 个基站，每个基站的信息都可能被手机收到，因为它们使用的输出信道都是 $\bar{a}$。假设收到的是基站 P_1 发来的信息 b，“存于”z 中（此时 R 的两个名字 z 都变成名字 b）。R 就可以通过第一个 $\bar{b}$ 输出信道把通信目标的地址 f 传给 P_1，“存于”c 中（此时名字 c 变成名字 f）。然后手机使用者通过信道 m 输入短信内容 g，“存于”n 中（此时名字 n 变成名字 g）。R 可以通过第二个 $\bar{b}$ 信道向 P_1 输出短信 g，存于 u 中（此时名字 u 变成名字 g）。又由于 P_1 的信道 c 已被置换成 f，基站 P_1 可以把短信 g 发往目的地。这个过程用变迁公式表示也许更加清楚，其中 $P_{2-n} \overset{\text{def}}{=} P_2 \mid \cdots \mid P_n$：

$$P = P_1 \mid P_{2-n} \mid R \xrightarrow{xf} P_1 \mid P_{2-n} \mid a(z).\,\bar{z}f.\,m(n).\,\bar{z}n.\,R$$

$$\xrightarrow{\tau} b(c).\,b(u).\,\bar{c}u.\,P_1 \mid P_{2-n} \mid \bar{b}f.\,m(n).\,\bar{b}n.\,R \xrightarrow{\tau} b(u).\,\bar{f}u.\,P_1 \mid P_{2-n} \mid m(n).\,\bar{b}n.\,R$$

$$\xrightarrow{mg} b(u).\,\bar{f}u.\,P_1 \mid P_{2-n} \mid \bar{b}g.\,R \xrightarrow{\tau} \bar{f}g.\,P_1 \mid P_{2-n} \mid R \xrightarrow{\bar{f}g} P_1 \mid P_{2-n} \mid R$$

这里有 n 个基站，理论上，每个基站都可以和手机通信。这里仅给出了手机 R 和基站 P_1 通信的情况。理论上，和其他任意一个基站 P_i 的通信都是可能的。在现实中有很多约束。在同等功率的情况下，最主要的是看信号传递距离。离手机远的基站根本接收不到手机发出的信号。一般情况下与双方的信号强度有

关，并且随着手机的移动，能够和手机通信的基站也在变化。实际的通信远比这个例子复杂得多。例如这里就没有考虑蜂窝区切换，还可以从例子中挑出很多不符合实际的瑕疵，但是这个例子已足以说明在 π 演算中信道名可以被传输的好处。还有一点需要说明：在上面列出的变迁序列中 P_1 和 R 都是递归调用，因此所有的输入都不会影响到递归程序体中的名字。 ▮

现在看第二个特色：π 演算的屏蔽适用于一切名字吗？回答是原则上适用于一切名字，无论它是在信道名的岗位上，还是在被传输的值的岗位上。但是对信道名岗位上的名字的屏蔽是绝对的，而对被传输值的名字的屏蔽则是相对的，有条件的，在某些情况下可以从宽处理。在 π 演算中，有三个变迁 EXTRU，INTRU 和 OPEN（见 7.2 节）就是说明在什么情况下可以从宽处理的。而这三个变迁又都涉及输出。π 演算有两类输出，一类是自由输出 $\bar{a}z$，相当于 CCS 的输出；另一类是约束输出 $\bar{a}(z)$，在 CCS 中是没有的。先说自由输出 $\bar{a}z$。它由变迁规则 INTRU 规定。在这里，无论是信道名 $\bar{a}$ 还是输出的值 z，在输出时都没有受到屏蔽，故而称之为自由的。但是输入方动程虽不允许有 (νa) 的帽子，却允许有 (νz) 的帽子。这表示若对方有信道名 (νa) 的屏障，则通信不能进行。而值的屏障 (νz) 却是可以穿越的，为此只需满足一个条件：换名。本来，带屏蔽的动程 $(\nu z)P$ 是不可以从外界接收名字 z 的。但是考虑到 z 是约束变量，可以事先用 α 转换把它“改名换姓”，然后就可以从外界接收名字 z 了。于是变迁 $\bar{a}z.P \mid (\nu z)a(y).Q \xrightarrow{\tau} P \mid (\nu w)Q\{\alpha: w/z\}\{z/y\}$ 就成为合法的了，即使把 $\bar{a}z.P$ 换成 $(\nu t)\bar{a}z.P$ 也是一样。当然前提是 w 对于 Q 来说是一个完全新的名字。这个通信侵入了 (νw) 的作用域，因而称之为作用域入侵规则，所以用 INTRU 命名。现在来说约束输出 $\bar{a}(z)$。$\bar{a}(z)$ 不是动程前缀，而是一个变迁动作，它相当于执行动作 $(\nu z)\bar{a}z$，可以理解为：自由输出 $\bar{a}z$ 突破值的屏蔽 (νz) 等于约束输出 $\bar{a}(z)$。它首先涉及 OPEN 规则，该规则规定在 $(\nu z)P$ 形式的动程中，P 执行输出动作 $\bar{a}z$ 并非完全不可能，只是要加上一个条件：执行输出动作 $\bar{a}z$ 时要把前面的“帽子”(νz) 一起扔掉。具体说来是：在自由输出 $P \xrightarrow{\bar{a}z} P'$ 成立的前提下，约束输出 $(\nu z)P \xrightarrow{\bar{a}(z)} P'$ 是合法的，P 的帽子 (νz) 被脱掉了。当然帽子不能随便地消失，那么帽子去了哪里呢？考查另一个操作 EXTRU 便见分晓。该操作的规则是，如果加上输入操作 $Q \xrightarrow{az} Q'$，并且 a 不是 Q 的自由名字，则 $(\nu z)P \mid Q \xrightarrow{\tau} (\nu z)(P' \mid Q')$ 是一个合法变迁，原来只戴在 P 头上的帽子现在戴到了 P'、Q' 双方的头上。这是一次完美的通信操作，不过不是由自由输出，而是由约束输出完成的。在这里最关键的是，输出方不仅输出了名字 z，而且还输出了屏蔽 (νz)。就是说输出