《领域特定语言》

图书基本信息

书名：《领域特定语言》

13位ISBN编号：9787111413059

10位ISBN编号：7111413059

出版时间：2013-3

出版社：机械工业出版社华章公司

作者：Martin Fowler

页数：488

译者：ThoughtWorks中国

《领域特定语言》

前言

在我开始编程之前，DSL（Domain – Specific Language，领域特定语言）就已经成了程序世界中的一员。随便找个UNIX或者Lisp老手问问，他一定会跟你滔滔不绝地谈起DSL是怎么成为他的镇宅之宝的，直到你被烦得痛不欲生为止。但即便这样，DSL却从未成为计算领域的一大亮点。大多数人都是从别人那里学到DSL，而且只学到了有限的几种技术。我写这本书就是为了改变这个现状。我希望通过本书介绍的大量DSL技术，让你有足够的信息来做出决策：是否在工作中使用DSL，以及选择哪一种DSL技术。造成DSL流行的原因有很多，我只着重强调两点：首先， 提升开发人员的生产力；其次，增进与领域专家之间的沟通。如果DSL选择得当，就可以使一段复杂的代码变得清晰易懂，在使用这段代码时提高程序员的工作效率。同时，如果DSL选择得当，就可以使一段普通的文字既可以当做可执行的软件，又可以充当功能描述，让领域专家能理解他们的想法是如何在系统中得到体现的，开发者和领域专家的沟通也会更加顺畅。增进沟通比起工作效率提升困难了一些，但带来的效果却更为显著。因为它可以帮助我们打通软件开发中最狭窄的瓶颈—程序员和客户之间的沟通。我不会片面夸大DSL的价值。我常常说，无论你什么时候谈到DSL的优缺点，你都可以考虑把“ DSL ”换成“ 库 ”。实 际上，大多数DSL都只是在一个框架或者库上又加了薄薄的一层外壳。于是，DSL的成本和收益往往会比人们预想的要小，但也未曾得到过充分的认识。掌握良好的技术可以大大降低构造DSL的成本，我希望这本书可以帮你做到这一点。这层外壳虽薄，却也实用，值得一试。为什么现在写这本书DSL已问世很长时间，但近些年来，它们掀起了一股流行风潮。与此同时，我决定用几年的时间写这本书。为什么呢？虽然我并不知道自己能否给这股风潮下一个权威的定义，但是可以分享一下自己的观点。在千禧年到来的时候，编程语言界—至少在我的企业软件世界里—隐约出现了一种颇具统治性的标准。先是 Java，它在几年的时间里风光无限。即使后来微软推出的C#挑战了Java的统治地位，但这个新生者依然是一门跟Java很相似的语言。新时代的软件开发被编译型的、静态的、面向对象的、语法格式跟C类似的语言统治着（甚至连VB都被迫变得尽可能具有这些性质）。然而人们很快发现，并不是所有的事情都能在Java/C#的霸权下运作良好。有些重要的逻辑用这些语言无法很好实现，这导致了XML配置文件的兴起。不久之后，有程序员开玩笑说，他们写的XML代码比Java/C#代码都多。这固然有一部分原因是想在运行时改变系统行为，但也体现出人们的另外一个想法：用更容易定制的方式表达系统行为的各个方面。虽然XML噪音很多，但确实可以让你定义自己的词汇，而且提供了非常强大的层次结构。不过后来人们实在无法忍受XML那么多的噪音了。他们抱怨尖括号刺伤了他们的双眼。他们希望一方面能够享受XML 配置文件带来的好处，另一方面又不用承受XML的痛苦。我们的故事到现在讲了一半，这个时候Ruby on Rails横空出世，耀眼的光芒让一切都褪尽了颜色。无论Rails这个实用平台在历史上会占据什么样的位置（我觉得这确实是个优秀的平台），它都已经给人们对于框架和库的认识带来了深远的影响。Ruby社区有一个很重要的做事方式：让一切显得更加连贯。换句话说，在调用库的时候，就像用一门专门的语言进行编程一样。这不禁让我们回想起一门古老的编程语言：Lisp。通过它我们也看到了Java/C#这片坚硬的土地上绽开的花朵：连贯接口（fluent interface）在这两门语言中都变得流行起来，这大概要归功于JMock和 Hamcrest的创始人的不断努力。回头看看这一切，我发现这里面有知识壁垒。有的时候，使用定制的语法会更容易理解，实现也不难，人们却用了XML；有的时候，使用定制的语法会简单很多，人们却把Ruby用得乱七八糟；有的时候，本来在他们常用的语言中使用连贯接口就可以轻易实现的事情，人们非要使用解析器。我觉得这些事情都是因为存在知识壁垒才发生的。经验丰富的程序员对DSL的相关技术所知寥寥，没法对使用哪一项技术做出明智的判断。我对打破这个壁垒很感兴趣。为什么DSL很重要本书2.2节会讲述更多细节，不过我觉得需要学习DSL（以及本书中提到的其他技术）的原因主要有两个。第一个原因是提升程序员的生产力。先看下面这段代码：input =~ /\d{3}-\d{3}-\d{4}/你会认出这是个正则表达式匹配，也许你还知道它匹配的是什么。正则表达式常常由于过于费解而遭受指责，但试想一下，如果你所能够使用的都是普通的正则控制代码，这段模式匹配会变成什么样子。而这段代码跟正则表达式相比，又是何等容易理解，容易修改？DSL的第一个优势是它擅长在程序中的某些特定地方发挥作用，并且让它们容易理解，进而提高编写、维护的速度，也会减少bug。DSL的第二个优势就不仅仅限于程序员的范畴了。因为DSL往往短小易读，所以非程序员也能看懂这些驱动他们重要业务的代码。把这些真实的代码暴露在理解该领域的人们面前，可以确保程序员和客户之间有非常顺畅的沟通渠道。当人们谈论这类事情的时候，他们常说DSL可以让你不再需要程序员。我对这一论调极度不认同，毕竟那是说COBOL的。

不过也确实有些语言是由那些自称不是程序员的人来用的，比如CSS。对这种语言来说，读比写要重要得多。如果一个领域专家可以阅读并且理解核心业务代码中的绝大部分，那他就可以跟写这段代码的程序员进行深入细节的交流。第二个原因是使用DSL并非易事，不过回报也是相当丰厚的。软件开发中最狭窄的瓶颈就是程序员和客户之间的沟通，任何可以解决这一问题的技术都值得学习。别畏惧这本大厚书看到这本书这么厚，你可能会吓一跳；我自己发现要写这么多内容的时候都忍不住倒吸一口冷气。我对大厚书的态度总是小心翼翼，因为我们用来阅读的时间是有限的，一本厚书就意味着时间上的大量投资。因此在这种情况下我更倾向于使用“姊妹篇”的方式。姊妹篇实际上是关于一个主题的两本书。第一本是叙述性质的书，需要仔细阅读。我希望它可以大致地描述出这个主题的主要内容，让读者有个整体认识就好，不用深入细节。我觉得叙述部分最好不要超过150页，这是个比较合理的厚度。第二本书是参考资料，不需要一页一页翻阅（虽然有些人也这样看）。用的时候再仔细看就行。有些人喜欢先读完第一本，有了整体认识之后，再去看第二本书里面感兴趣的章节。有些人喜欢一边读第一本，一边找第二本中感兴趣的地方读。我之所以采用这种划分方式，主要还是想让你们了解哪些地方可以跳过，哪些地方不能忽略，这样你也就可以有选择地深入阅读了。我已经尽力让参考资料那部分独立成篇了，如果你想让某人使用“树的构建”（第24章），就让他阅读那个模式，即使他可能对DSL没有清楚的认识，但是也能知道怎么做。这样一来，一旦你完全理解了概述部分，这本书就变成了参考资料，想查详细资料的话，一翻开就能找到。本书之所以这么厚，是因为我没能找到把它变薄的方法。它的一个主要目的是分析、比较DSL的各项技术。讨论代码生成、Ruby元编程、“解析器生成器”（第23章）工具的书有很多，本书涵盖所有这些技术，让你可以了解它们的异同。它们都在广阔的舞台上扮演着自己的角色，在帮助你了解这些技术之外，我还想介绍一下这个舞台。本书主要内容本书旨在全面介绍各种DSL及其构造方式。当人们尝试DSL的时候，经常就只选一种技术。你可以在本书里看到对多种技术的介绍，真正用的时候就可以做出最合适的选择。本书还提供了很多DSL技术的实现细节和例子。当然，我无法把所有的细节都写下来，但也足以使读者入门，在早期决策时起到辅助作用。前3章讲述什么是DSL、DSL的用途以及DSL与框架和库的区别。第5章和第6章可以帮你理解如何构建外部DSL和内部DSL。第三部分讲述解析器所扮演的角色，“解析器生成器”（第23章）的作用，用解析器解析外部DSL的各种方式。第四部分展示了在一种DSL风格中所能使用的多种语言结构。虽然它不能告诉你怎样充分利用你钟爱的语言，却能帮助你理解一门语言中的技术在不同语言之间的对应关系。第五部分介绍其他计算模型，有助于读者学习如何构建模型。第六部分列出了生成代码的各种策略，你需要的话可以看一下。第9章简单介绍了新一代的工具。本书所介绍的绝大部分技术都已经面世很长时间了；语言工作台更像是未来科技，虽然应当有美好的前景，但没有经验证明。本书读者对象本书的理想读者是那些正在思考构建DSL的职业软件程序员。我觉得这种类型的读者都应该有多年的工作经验，认同软件设计的基本思想。如果你深入研究过语言设计的话题，那这本书里大概不会有什么你没有接触过的资料。我倒是希望我在书中整理并表述信息的方式对你有所帮助。虽然人们在语言设计方面做了大量的工作—尤其是在学术领域，可这些成果能够为专业编程领域服务的却寥寥无几。叙述部分的前几章还可以澄清一些困惑，比如什么是DSL，DSL有什么用途。通读第一部分以后，你就可以更全面地掌握DSL的不同实现技术。这是本Java书或者C#书吗本书和我曾写过的大部分书一样，与编程语言没有多大关系。我最想做的事情是揭示一些与编程语言无关的通用原则和模式。因此，不管你用的是哪种流行的面向对象语言，本书里的思想都会为你提供帮助。函数式语言可能会是一条代沟。虽然我觉得很多内容依然对函数式语言适用，但我在函数式编程中的经验并不足以让我做出判断，它们的编程范式到底会从多大程度上影响到书中的建议。本书对于过程式语言（即非面向对象的语言，例如C）的作用也很有限，因为我讲的很多技术都依赖于面向对象技术。虽然我写的是通用原则，但为了能够把它们恰当地讲述出来，我还需要一些例子—于是需要一门具体的编程语言。在选择用哪门语言来写例子的时候，我的首要标准是有多少人能读懂它。于是绝大多数例子都是用Java或C#写的。这两门语言在业界广泛使用，有很多相似之处：类C的语法、内存管理，为人们提供各种便利的类库。但我的意思可不是说它们就是写DSL的最佳选择了（这里需要特别强调一下，因为我根本就不认为它们是最佳选择），只是说它们最能够帮助读者理解我讲的通用概念。我尽力让二者出现的机会均等，只有当使用某种语言更方便的时候，我的天平才会稍稍倾斜一下。虽然内部DSL的良好运用常常要用到某些另类的语法特色，但我也尽力避免使用一些需要太多语法知识才能理解的语言元素，着实挺困难的。还有一些思想是必须使用动态语言才能满足的，没法用Java或C#实现。在那些情况下我就换用Ruby，因为这是我最熟悉的动态语言。它为我提供了很多帮助，因

为它的特性完美地契合了编写DSL的需求。另外再强调一点，虽然我个人更熟悉某种语言，在选择时也考虑了个人偏好，但这并不能推断出这些技术换个地方就不能用了。我很喜欢Ruby，但如果你想看看我对语言的偏执，那只有贬低Smalltalk才行。值得一提的是，另外有许多语言都适合构建DSL，尤其还有一些是专门为了写内部DSL而设计出来的。我之所以没有提到它们，是因为我对它们所知不多，没有足够的信心进行评价。你不要认为我对它们有什么负面观点。另外还要提一句，在写这本书的时候，本来试图和语言无关，可大多数技术的实用性偏偏都要直接依赖于某种语言的特性。这是让我最苦恼的地方了。我为了实现大范围内的通用性做出了很多权衡，但你必须意识到，这些权衡可能会因具体的语言环境而彻底改变。本书缺少什么在写这样一本书的过程中，要说什么时候最让人垂头丧气、濒临崩溃，莫过于自己意识到必须停笔的那一刻了。我为这本书花费了几年的时间，我相信这里面有很多值得你阅读的内容。但我也知道我留了很多疏漏之处。我本来想弥补这些疏漏，可这得占用大量时间。我的信念是，宁可出版一本未完成的书，也不要再等上几年把书写完—即便是真的能够写完的话。下面简单介绍一下我实在没时间补充的内容。前面曾提到过一点—函数式语言。实际上，在当代那些基于ML和/或Haskell的函数式语言中，构建DSL已经是广为人知的事实了。我在书中基本没提这部分内容。我也很想知道，当我熟悉了函数式语言及其DSL应用之后，本书的内容安排会发生多大的改变。有一件事情最让我心神不安，那就是没有把近期有关诊断和错误处理的讨论加进去。我记得上大学的时候曾学到过，诊断这件事情在写编译器的过程中是何等艰难，因此忽略这一点让我觉得自己像是在作表面文章。我个人最喜欢第7章。我本来可以写很多内容的，只可惜时不我予。最后我只好决定少写一些—希望它依然可以激发你主动探索更多模型的兴趣。章节引用虽然本书的结构比较普通，但引用章节的结构还是需要稍稍介绍一下的。我把引用章节分成一系列主题，按照相似性组成不同的章节。我的想法是每个主题都可以独立成篇，于是你读完第一部分以后，就可以任选一个主题深入了解，无须再涉及其他章节。如果有例外情况的话，我会在对应主题的开篇提到。大部分主题都以模式的形式呈现。模式是对于一再重复出现的问题的通用解决方案。所以如果你有一个常见的问题：“我该怎么处理我的解析器结构呢？”对这个问题的两种可行模式是“分隔符指导翻译”（第17章）和“语法指导翻译”（第18章）。在过去的二十年间，人们写了很多关于软件开发模式的书，不同的人有不同的视角。我的看法是，模式给我提供了一种组织参考资料的良好方式。第一部分告诉你如果想要解析文本，可以考虑上面两种模式。模式本身提供了更多的信息以供选择和具体实施。引用章节大都是以模式的结构来写的，但也有些例外：并不是所有的主题在我眼中都是解决方案。比如“嵌套的运算符表达式”（第29章），它的重点就不是解决方案，也不符合模式的结构，因此我没采用模式风格的描述方式。还有一些情况很难称为模式，比如“宏”（第15章）和“BNF”（第19章），可是用模式结构来描述它们却很合适。总的来说，只要是模式结构—尤其是把“如何起作用”和“何时使用模式”分离开这种形式—能够帮我描述概念，我就一直在使用它。模式结构大多数作者在写模式的时候都用了一些标准模板。我也不例外，既用了一个标准模板，又与别人用的不一样。我所用的模板，或称模式形态，是我第一次用在企业应用架构模式（P of EAA，[Fowler PoEAA]）中的模式。它的形式如下。模板中最重要的元素大概要数名字。我喜欢用模式来描述各个引用主题，最大的原因就是它可以帮我创建一个强大的词汇表，方便展开讨论。虽然这个词汇表不一定能得到广泛应用，但至少可以让我的写作保持一致性，也可以在别人想要用这个模式的时候，给他提供一个起始点。接下来的两个元素是意图和概要。它们对模式进行简要的概括。它们还能起到提醒作用，如果你已“将模式纳入囊中”，但忘了名字，它们可以轻轻拨动你的记忆。意图用一句话总结模式，而概要是模式的一种可视化表示—有时候是一张草图，有时候是代码示例，不管是什么形式，只要能够快速解释模式的本质就好。我有时候使用图表，有时候会用UML画图，不过要是有其他方式可以更容易表达意图，我也是很乐意用的。接下来就是稍长一些的摘要了。我一般会在这一部分给出一个例子，用来说明模式的用途。摘要由几段话组成，同样是为了让读者在深入细节之前先了解模式全貌。模式有两个主体部分：工作原理和使用场景。这两部分没有固定顺序，如果你想了解是否该用某个模式，可能就只想读“使用场景”这一节。不过，一般来说，不了解工作原理的话，只看“使用场景”是没什么作用的。最后一部分是例子。我尽力在“工作原理”这一节把模式的工作原理讲清楚，但人们一般还是需要通过代码来理解。然而，代码示例是有危险的，它们演示的只是模式的一种应用场景，可有些人却会以为模式就是这个用法，没有理解背后的概念。你可以把一种模式用上千百遍，每次稍稍有些差异，可我没有地方容纳那么多代码，所以请记住，模式的含义远远不止从特定示例中看到的那么多。所有的例子都设计得非常简单，只关注要讨论的模式本身。我用的例子都是相互独立的，因为我的目标是每一个

引用章节都独立成篇。一般来说，在实际应用模式的时候还需要处理其他大量问题，但在一个简单的例子中，起码可以让你有机会理解问题的核心。丰富的例子更贴近现实，可它们也会引入大量与当前模式无关的问题。于是我只会展示一些片段，你需要自己把它们组装起来，满足特定的需求。这同时也意味着我在代码中主要追求的是可读性。我没有考虑性能、错误处理等因素，它们只会把你的注意力从模式的核心转移到别处。也基于这个原因，我力图避免写一些我觉得很难借鉴的代码，即便是更符合该语言的惯例也不予考虑。这个折衷在内部DSL上会显得有些笨拙，因为内部DSL经常要靠语言的小窍门来强化语言的连贯性。很多模式都缺少上面讲的一两个部分，因为确实没什么可写的。有些模式没有例子，因为最合适的例子在其他模式里面用到了—在发生这种情况的时候，我会把它指出来。致谢当我每次写书的时候，很多人都为我提供了大量帮助。虽然作者那里写的是我的名字，但许多朋友都为提升本书的质量作出了巨大的贡献。我首先要感谢的是我的同事Rebecca Parsons。我对DSL这个主题曾有很多顾虑，例如，它会涉及很多学术背景的知识，而那些是我所不熟悉的。Rebecca有深厚的语言理论背景，她在这方面给了我很多帮助。此外，她也是我们公司的首席技术探路人和战略家，她可以将她的学术背景和大量的实践经验合二为一。她有能力，并且也愿意为本书付出更多心血，但ThoughtWorks在其他方面更需要她。我很高兴，我们曾关于DSL这个主题滔滔不绝。作者总是希望（并且带着小小的恐惧）审校者可以通读全书，找出不计其数的大大小小的错误。我幸运地找到了Michael Hunger，他的审校工作做得极其出色。从本书刚刚出现在我网站上的时候，他就开始不断地给我挑错，催促我改正—这正是我需要的态度。他同时也催促我详细介绍使用静态类型的技术，尤其是静态类型的“符号表”（第12章）。他给我提供了无数建议，足以再写两本书了。我希望有朝一日可以把这些想法写下来。在过去的几年里，我和同事们包括Rebecca Parsons、Neal Ford、Ola Bini，写过很多这方面的文章。在本书里，我也借鉴了他们的一些成型的想法。ThoughtWorks慷慨地给予我大量时间来写这本书。在花了很长时间决定不再为某一家公司工作之后，我很高兴找到一家让我愿意留下来，并且积极地参与其建设的公司。本书还有很多正式的审校者，他们为本书提供了大量建议，找出了很多错误。他们是：David Bock David IngGilad Bracha Jeremy MillerAino Corry Ravi MohanSven Efftinge Terance ParrEric Evans Nat PryceJay Fields Chris SellsSteve Freeman Nathaniel SchuttaBrian Goetz Craig TavernerSteve Hayes Dave ThomasClifford Heath Glenn VanderburgMichael Hunger我还要感谢David Ing，是他提出了第10章的章名。成为一个系列丛书的编辑之后，我就有了些美妙的特权，比如，我拥有了一个很出色的作者团队，他们可以帮我出谋划策。其中我尤其要感谢Elliotte Rusty Harold，他提供了很多精彩的建议和评论。我在ThoughtWorks的很多同事也成为我想法的源泉。我要谢谢过去几年里允许我在项目中闲逛的每个人。我能写下来的远不及所看到的，能在这样广袤的海洋中徜徉，我感到无比愉悦。有些人给Safari联机丛书的初稿提供了很多建议，我在正式付梓之前也参考了他们的想法。这些人是：Pavel Bernhauser、Mocky、Roman Yakovenko、tdyer。我还要谢谢本书出版商Pearson的工作人员。Greg Doench是本书的组稿编辑，他负责出版的整体流程。John Fuller是本书的执行编辑，他监管生产流程。Dmitry Kirsanov调整了我拙劣的英语，让本书语言流畅。Alina Kirsanova组织了本书的布局。

《领域特定语言》

内容概要

本书是DSL领域的丰碑之作，由世界级软件开发大师和软件开发“教父”Martin Fowler历时多年写作而成，ThoughtWorks中国翻译。全面详尽地讲解了各种DSL及其构造方式，揭示了与编程语言无关的通用原则和模式，阐释了如何通过DSL有效提高开发人员的生产力以及增进与领域专家的有效沟通，能为开发人员选择和使用DSL提供有效的决策依据和指导方法。
全书共57章，分为六个部分：第一部分介绍了什么是DSL，DSL的用途，如何实现外部DS和内部DSL，如何生成代码，语言工作台的使用方法；第二部分介绍了各种DSL，分别讲述了语义模型、符号表、语境变量、构造型生成器、宏和通知的工作原理和使用场景；第三部分分别揭示分隔符指导翻译、语法指导翻译、BNF、易于正则表达式表的词法分析器、递归下降法词法分析器、解析器组合子、解析器生成器、树的构建、嵌入式语法翻译、内嵌解释器、外加代码等；第四部分介绍了表达式生成器、函数序列、嵌套函数、方法级联、对象范围、闭包、嵌套闭包、标注、解析数操作、类符号表、文本润色、字面量扩展的工作原理和使用场景；第五部分介绍了适应性模型、决策表、依赖网络、产生式规则系统、状态机等计算模型的工作原理和使用场景；第六部分介绍了基于转换器的代码生成、模板化的生成器、嵌入助手、基于模型的代码生成、无视模型的代码生成和代沟等内容。

《领域特定语言》

作者简介

Martin Fowler，世界级软件开发大师，软件开发“教父”，敏捷开发方法的创始人之一，在面向对象分析与设计、UML、模式、极限编程、重构和DSL等领域都有非常深入的研究并为软件开发行业做出了卓越贡献。他乐于分享，撰写了《企业应用架构模式》（荣获第13届Jolt生产力大奖）、《重构：改善既有代码的设计》、《分析模式：可复用的对象模型》、《UML精粹：标准对象建模语言简明指南》等在软件开发领域颇负盛名的著作。

《领域特定语言》

书籍目录

《领域特定语言》

章节摘录

译者序2008年，老马（Martin Fowler）在Agile China上做主旨发言，题目就是领域特定语言（Domain Specific Language，DSL）。老马提携后辈，愿意跟我合作完成这个演讲。而我呢，一方面，年少轻狂认为这个领域我也算个中好手，另一方面，也感激老马的信任和厚爱，就答应了。当时我已经知道老马在写一本关于这个主题的书，便 跟他讨要原文来看。当时还没有成型的稿子，只有非常简略的草稿和博客片段。 2010年年底，ThoughtWorks技术战略委员会（ThoughtWorks Technology Advisor Board）在芝加哥开会。那时候，这本书的英文原版已然出版。晚上聚餐的时候，我心血来潮，跟老马说如果有机会，希望能将这本书翻译成中文，介绍给中国的开发者。老马听了很高兴，把最终定稿的电子版访问权限授予了我。几个月后，同事刀哥（李剑）问我有没有兴趣参加这本书的翻译，我说当然有。后来回想起来，我还是上了刀哥的当，因为突然之间我就从参与翻译变成负责翻译了。由于我手里已经有原稿的缘故，因此我们没有采用出版社提供的Word版本，而是在英文原稿上直接翻译。原稿除了文字部分之外，还有几千行代码。其中包括XSTL、Ruby、C++、Java、图像处理脚本，甚至还有用于构建样书的rake脚本。惊讶之余，作为程序员的求知欲也被激发出来了。在动手翻译之前，我们花了两天彻底了解这些代 码的作用。然后就发现了这个惊人的秘密：这是一本用领域特定语言写就的关于领域特定语言的书。原文的文字部分是老马在docbook基础上定义的领域特定语言。这种语言除了在docbook的基础结构上定义了章节模板之外，还有两个专用结构：patternRef和codeRef 。patternRef用于处理模式名称在不同章节中的引用。本书分为两部分，前面的概念讲述部分和后面的模式部分。它耗时3年写就，在草稿阶段模式名称都未确定，各章节之间交叉引用很多。一旦出现模式名称改变，更新同步成本就很高。为此，老马定义了专有的语言结构，patternRef。所有对于模式的引用，都通过patternRef实现。由 patternRef解析处理应该使用那个具体的名称。这个巧妙的做法在后来的翻译中给我们带来了很大的困扰。因为patternRef会处理英语中的单复数，而中文不会有这样的情况。翻译稿中出现了大量的s和es。最后还是通过修改DSL解析器里才解决了这个问题。codeRef则表示代码引用。这本书属于技术领域，其中会有大量代码示例；同一份代码示例会在不同章节中引用，一旦写法变化，就需要同步检查它在上下文内是否还能起到示范作用。老马先在示范代码的源代码中通过注释加入XML标注，把代码分解成一段段可引用的例子。因为是代码注释，所以不会影响源代码的编译、调试和重构。 然后，再通过codeRef，表明是哪个例子的哪段示例。最后，再通过Ruby和XSLT，摘取对应的代码段，生成相应的 文本。我一直认为在澄清概念和发现模式上老马是有超能力的。通常会忘记他也是个ThoughtWorker，而让做事情变 得有趣，则是每个ThoughtWorker都有的超能力。翻译这本书并不轻松，其中很多概念中文并无定译。为了呈现最好的结果，我们成立了一个翻译小组，包括熊节、郑烨、李剑、张凯峰、金明等有较多翻译经验的ThoughtWorker悉数在内。虽然如此，仍然难免疏漏，望读者 不吝斧正。徐昊　　ThoughtWorks中国

《领域特定语言》

编辑推荐

《华章程序员书库:领域特定语言》全面详尽地讲解各种DSL及其构造方式，揭示与编程语言无关的通用原则和模式，阐释如何通过DSL有效提高开发人员的生产力以及增进与领域专家的有效沟通。

精彩短评

《领域特定语言》

精彩书评

1、DSL领域丰碑之作,世界级软件开发大师和软件开发“ 教父 ”Martin Fowler的著作， thoughtworks中国翻译。全面详尽地讲解了各种dsl及其构造方式，揭示了与编程语言无关的通用原则和模式，阐释了如何通过dsl有效提高开发人员的生产力以及增进与领域专家的有效沟通，能为开发人员选择和使用dsl提供有效的决策依据和指导方法。真没想到这本书快要出来了，等了三年了，说什么都要买一本，一个被忽视了N年的编辑利器。
2、讲内部DSL的部分 还行 挺不错的， 但是没学到什么 或者只是对之前的有些想法产生了共鸣，不知道是不是翻译问题，反正这本书的例子感觉很不好。而且通篇几乎都是基于一个例子扩展开来，很生涩。 外部DSL的部分，不如去自己嚼一下龙叔，或者那本简单的编程语言实现模式。 章节分太细，每章连皮毛都不是，只能就是个概略
3、看到评价才买的，买来粗读了一遍，实在比较失望。首先，书的翻译水平欠佳。其次，内容组织缺乏逻辑性，缺乏对领域定义语言的明确定义、分类、处理方法的完整、一致的介绍。感觉作者写作时东拉西扯，虽然也算言之成理，但是并没有看到什么真知灼见。第三，各种模式缺乏足够的背景介绍和应用说明，并且对大多数模式的介绍都浅尝辄止。总之，这本书没有相关背景知识的人读起来会觉得不知所云，有相关背景的人读起来可能也会觉得有点乏味，当做手册用觉得太单薄，当教材用觉得太凌乱，当谈资用觉得太生硬，我还是拿它当枕头算了。
4、书是 130325 寄送到珠海的,超乎相象的厚!- 好象趁出差帝都,用原先的家务时间替换为读书时间,才将此书啃完的- 果断是当初: 编程语言实现模式 (豆瓣)http://book.douban.com/subject/10482195/的配套图书! 之前就吼过:"但是,要配套另外两本才真心实用的起来,这本是基本地图,ANTLR 是装备,DSL 语法设计图书则是藏宝图了..."照例吐槽先!- 至今读到的最象 OREILLY 图书的技术图书!- 不论书脊,排版,章节设计- 都是浓浓的一般 OREILLY 动物书的自在流畅- 山寨到这种地步真心不易- 但是,太对不起 ThoughtWorks 的敏捷形象了吧- 以外,最最最怨恨的是居然没有给出图书源代码工程来!- 最最最吸引俺的,就是译序中指出的:- "这是用DSL写成的有关DSL的书"!- 因为OBP 工程的原因http://code.google.com/p/openbookproject/- 俺是习惯使用 rST 配合 Sphinx 进行图书撰写的- 也同样遇到了老马的 patternRef/codeRef 问题- 虽然 rST 本身提供了一定的解决方案,但是,用起来还是不够直觉- 如果 DSL 可以公开 docbook 的工程源代码- 那么,真心可以是对中国出版社技术图书组稿方式的一大推动- 其实不用全部公开,只对中/E文版本公开包含图书典型场景的几章就好!整本图书,正如作者在序言中所述- "尽力将DSL 领域所有知识,进行了关键性评点..."- 就已经这么厚了! 所以,作者只能自认没法真正完成本身就在快速迭进的DSL技术全貌描述- 所以,整本内容,处处表现出一种谦虚谨慎的学术范儿,令人安心,共鸣- 只是,出版社应该联系作者,尽快将DSL 实现过程把诊断/错误处理相关的内容,专门出本书吧!其实,对于多数程序猿,第一部分,叙述部分的内容,就已经值得整书价格了!- 对什么是DSL,DSL 的分类,DSL 的适用场景,进行了综合性介绍- 可以感觉的出,是作者写的最畅快的部分- 然后的部分:通用/外部DSL/内部DSL/其它模型,都是具体实现的细节论述- 整个图书的设计,非常象当年的 可爱的Pythonhttp://book.douban.com/subject/3884108/- 先用实例故事将想传达的经验包含进来- 通过 PCS 部分,对各个技术细节进行深入描述- 所以,读起来很是自得- 只是非常怨念的是作者展示各种DSL 技巧使用的是 JAVA/C#/Ruby - 几种俺最无爱的语言!- 特别是JAVA 看着大片大片密仄的代码段,实在没有心情认真理解- 好在,每种实现都不算长,也就一两页,当需要时,完全可以对等翻译成 py 版的- 其实,DSL 在平时就在经常使用- 各种框架使用的模板系统- ORM 库- OpenResty 的 lua 业务脚本- 复用 Graphviz 脚本,解析成工作流- 周蟒其实就是种经典的解析树操作实现的内部DSL- ...- 只是,没有机会真正自个儿从头实现一个- DSL 这书给出了几乎全部 DSL 实现形式- 对于各种场景什么适合,进行了郑重的建议- 省的我们自个儿乱折腾了- 完全能够根据业务特性,选择合适的模式,进行快速完成对于俺,DSL 这书最有用的是:- 第三部分,外部DSL,是个完备的指导手册,可以在具体实践中直接使用- 以及 p49 开始,不断强调的:- "对任何有持久价值的东西,都应该通过测试进行双重确认"- 这一思想,实在是软件工程师应该第一时间训练出来的自觉- 但是,,,,嗯嗯嗯,不展开了- 所以,要真心开始训练自个儿了...- 另外: ANTLR 看来是DSL 谁也跳不过去的工具,也应该有专门的图书进行分享吧?- 推荐 The Definitive ANTLR Reference: Building Domain-Specific Languages - Terence Parr 07年的书- 看哪个出版社手快了...PS:- 最后,忽然发现一个非常诡异的事儿:- 书是华章科技出的- 但是,在封底折边出现的是 博文的微信公众帐号宣传- 难道机械同电子出版社要合并了!?- 进一步的,博文的微信号名称是码农读书会- 嗯嗯嗯,图灵也乱入了?

5、这本书是不是太小众了，居然没有吐槽它的翻译？反正我最近是一边看一边备受折磨，忍不住要上来吐槽几句（处女吐槽）。已经糟到影响阅读的程度了！时不时要拿原版的出来对照，才理解是什么意思。姐姐，我就是英文烂，才要买中文版啊，能不能走点心，靠点谱啊!细节就不用说了，就挑几个标题来说（以下翻译意见不代表我的翻译就很好，只是说明中文版翻译得有多烂）：Context Variable 翻译成 语境变量，拜托，context在计算机领域一般是翻译成“ 上下文 ”，虽然这个也不怎么样，但至少约定俗成了，总比这个莫名其妙的“ 语境 ”好 Embedded Interpretation 翻译成 内嵌解释器，这字面哪有“ 器 ”啊，文章的意思也完全没有“ 器 ”的意思！而且你“ Embedded Translation ”翻译成“ 嵌入式语法翻译 ”，怎么 ” Embedded Interpretation " 就翻译成“ 内嵌解释器 ”了？ 上下两章的译法根本就不统一。Newline Separators 翻译成 以换行符作为分隔符，当然，这还算不上错，只是确实啰嗦。Delimiter-Directed Translation，Syntax-Directed Translation 分别翻译成“ 分隔符指导翻译 ”，“ 语法指导翻译 ”，这又是不可理解的中文，什么是“ 指导 ”啊！是灵道来指导工作吗？用“ 制导 ”虽然也不容易理解，至少还准确一些。最不能忍的是Alternative Computational Models翻译成其他计算模型概述？！！ 文章在此之前根本没提到“ 计算模型“ ，哪来 ” 其他计算模型“ ，而且哪来的 ” 概述“ ，这一部分全部都是概述好不，画蛇添足地加一个概述什么意思嘛！而且Alternative在这里根本就是“ 可选的 ”的意思好不！Method Chaining 翻译成 方法级联，这根本就不对，cascade 才是级联，这是“ 方法链"或 ” 方法链式调用“ Object Scoping 翻译成 ” 对象范围“ ，这不是“ 对象作用域“ 吗？Literal List 翻译成 列表的字面构造，“ 列表字面量 ”不就可以了？而Literal Map 则是直接选择不翻译！Transformer Generation翻译成 ” 基于转换器的代码生成“ ，Templated Generation 怎么就翻译成模板化的生成器了？能不能前后一致啊？Model-Aware Generation 翻译成 基于模板的代码生成 ，Model Ignorant Generation 翻译成 无视模型的代码生成了？一个是模型感知，一个是模型无知，感知不代表是“ 基于 ”，而“ 无知 ”是压根不知道，不是故意无视它标题尚且如此，内容就更不用说了，，很多句子完全就是机器翻译的水准！如果不是翻译们（对，是很多人合译的）水平太差，就是态度很有问题。

6、五星的主题，三星的内容，三星的行文，二星的翻译。前后文引用严重，概念混淆不清，语言啰嗦繁琐。懒得举例子了。

《领域特定语言》

章节试读

1、《领域特定语言》的笔记-第447页

误译一处，影响理解。

2、《领域特定语言》的笔记-第99页

误译一处，影响理解。
原文“动态”，译为“静态”。参见英文原文:
I find that Templated Generation works best when there's a lot of static code in the output and only a few dynamic bits—particularly since I can look at the template file and get a good sense of what gets generated.

3、《领域特定语言》的笔记-第399页

误译一处，影响理解。

4、《领域特定语言》的笔记-第237页

5、《领域特定语言》的笔记-第313页

书中代码给了一个区域安全检查的例子，C#，下面是我画的类图。

6、《领域特定语言》的笔记-第120页

语义模型与领域模型的不同:
A Semantic Model is a notion very similar to that of a Domain Model [poeaa]. I use a separate term because although Semantic Models are often subsets of Domain Models, they don't have to be. I use Domain Model to refer to a behaviorally rich object model, while a Semantic Model may be data alone. A Domain Model captures the core behavior of an application, while a Semantic Model may play a supporting role. A good example of this is an object-relational mapper that coordinates data between an object model and a relational database. You could use a DSL to describe object-relational mappings, and the resulting Semantic Model would consist of the Data Mappers [poeaa], not the Domain Model that is the subject of the mapping.

语言模型与抽象语法树的不同:
A Semantic Model is usually different from a syntax tree because they serve separate purposes. A syntax tree corresponds to the structure of the DSL scripts. Although an abstract syntax tree may simplify and somewhat reorganize the input data, it still takes fundamentally the same form. The Semantic Model, however, is based on what will be done with the information from a DSL script. It often will be a substantially different structure, and usually not a tree structure. There are occasions when an AST is an effective Semantic Model for a DSL, but these are the exception rather than the rule.

语言模型使得DSL不同于通用语言:
Traditional discussions of languages and parsing don't use a Semantic Model. This is part of the difference between working with DSLs and with general-purpose languages. A syntax tree usually makes a suitable structure to base code generation for a general-purpose language, so there's less desire to have a different Semantic Model. From time to time, a Semantic Model is used; for instance, a call graph representation is very useful for optimization.

Such models are referred to as intermediate representations—they are usually intermediate steps before code generation.

7、《领域特定语言》的笔记-第252页

foreign code
predicate
排除某些字符

8、《领域特定语言》的笔记-第208页

9、《领域特定语言》的笔记-第387页

C#语言实现类似这样的效果: 3.grams.flour

参考 Martin Fowler的代码，我做如下实现。
3.chi()，三尺 转换为国际标准单位制米。

```
namespace int_ext
{
  class Program
  {
    static void Main(string[] args)
    {
      System.Console.WriteLine( 3.chi() ); // output 0.999
    }
  }

  public static class int_e
  {
    public static double chi(this int arg)
    {
      return arg * 0.333;
    }
  }
}
```

10、《领域特定语言》的笔记-第448页

#define的高级用法，两处

11、《领域特定语言》的笔记-第259页

原文 Alternative Tokenization
意为 词法分析的变形，非常规的词法分析
嵌套引号的解析方法

转义
不常用的字符组合
多种分隔符供选择
成对花括号，利用 下推自动机的能力

12、《领域特定语言》的笔记-第188页

BNF中的递归。译文含义不清，不过正确的译法我也没想出来。

13、《领域特定语言》的笔记-第377页

益处
identifier
实现手段: 反射
以command为例

14、《领域特定语言》的笔记-第284页

如何使“ 无状态 ”的函数序列具有状态，使用语境变量。

15、《领域特定语言》的笔记-第225页

16、《领域特定语言》的笔记-第296页

Specification Pattern

17、《领域特定语言》的笔记-第290页

嵌套函数中，有多个同类型的参数，希望次序无关，如何区分这些参数。

问题
解决方案public enum Types { TopBorder, BottomBorder, LeftMargin, RightMargin, Transparent }

18、《领域特定语言》的笔记-第218页

参考ANTLR手册更好。

19、《领域特定语言》的笔记-第401页

lambda operator in C#.

20、《领域特定语言》的笔记-第106页

附图中的阴影被删除了，影响理解。

21、《领域特定语言》的笔记-第366页

子树与代码转换的能力
作者有趣的担心
递归遍历树
构建过程与讲解过程 的差异

22、《领域特定语言》的笔记-第243页

helper的位置
Semantic Model的容器
两个问题的解决方案:
分层语境，前向引用
hierarchic context and forward references.

23、《领域特定语言》的笔记-第211页

类图，继承和聚合关系，详图和简图；composite模式。

grammar用 parser comibinator in java描述
composite pattern
composite pattern应用于 parser comibinator，简图
继承和聚合关系，展开的详图

24、《领域特定语言》的笔记-第118页

误译一处，影响阅读。

25、《领域特定语言》的笔记-第279页

Expression Builder
单一Builder
多Builder

26、《领域特定语言》的笔记-第206页

Combinators are designed to be composed to create more complex operations of the same type as their input.

27、《领域特定语言》的笔记-第220页

红笔括起那段提到的不加EOF可能导致停止解析这个问题我从未遇到。哪位大侠了解，请指点我。先行感谢。

我又查了一下，Martin Fowler 在别的场合又提到一次这个问题，如下。

catalog : item* EOF;
...
Inevitably I did have problems even with this simple example. My biggest blocker was that I originally defined the catalog term as catalog : item*;, that is without the EOF. I then got confused because the parser didn't indicate an error when it got spurious input (like xitem foo). This wasn't helped by inconsistencies between Antlr and

AntlrWorks (the latter did show an error and older versions of AntlrWorks would handle whitespace differently too.)

[http://martinfowler.com/bliki/HelloAntlr.html]

28、《领域特定语言》的笔记-第445页

这一处，老马与我们的工作不同。
他使用更薄的DSL。

29、《领域特定语言》的笔记-第319页

各种语言称谓不同。
C语言可以用带有 void* 参数存储变量引用的 函数指针实现闭包。
问题的提出：用对象作为谓词，语法麻烦。
实例：C#2.0 & C#3.0
总结闭包：
1. 变量引用，而非复制，lexical scope；
2. lazy eval；
3. 闭包的创建、保存、执行。

《领域特定语言》

版权说明