

从“自动化”到“智能化”

——智能 Web 服务在信息处理中的应用

清华大学计算机系

杨文军 马路 丁峰 王克宏

目前无论是在当前工业界还是在学术界，Web 服务都被认为将导致下一代电子商务的革命。微软公司通过“一切都是服务”来概括 Web 服务将给当前 IT 业带来的冲击。

Web 服务的作用

在早期的分布式计算系统中，组件之间是紧密耦合的，不能有效地应用在互联网普遍存在的 B2B 电子商务上。因为组件之间的强耦合性要求系统遵循太多的来自于不同组织的协议和标准，而 B2B 电子商务系统是独立开发的，很难实现紧耦合的集成。未来的趋势是从紧耦合的单一系统发展为松耦合的动态绑定组件系统，Web 服务正是在这种趋势下出现的新的构造商务系统的技术手段。

Web 服务系统希望能使组件实现低耦合以及动态绑定。系统里的所有组件都是服务，它们封装了自己的行为并发布消息 API 给网络上其他参与协作的组件。不难看出，Web 服务反映了一种新的面向服务的方法，它带来的好处有以下一些：

- 实现及时 (just-in-time) 集成
- 减少封装的复杂度
- 能够集成遗留系统

目前 Web 服务在工业界和学术界分别沿着两个不同的方向进行研究。工业界注重于服务层的模块化，使之能够很快应用到商务系统中；而学术界注重于服务描述的表达。目前比较成熟的标准包括 SOAP、WSDL 和 UDDI。其中 SOAP(Simple Object Access Protocol)是在分散或分布式环境中交换信息协议，它基于；WSDL(Web Services Description Language)描述 Web 服务的接口以及服务的功能；UDDI(Universal Description, Discovery and Integration Service)是分布式 Web 服务的信息注册规范。除此之外，各大公司或组织也分别或联合发布了各种协议。

与工业界相比，学术界的目光要更长远一些。学术界主要致力于开发一种语言，能够把语义 Web 的研究成果引入到 Web 服务中，实现智能的 Web 服务。

智能 Web 服务

目前工业界推出的各种用 XML 语法来描述的 Web 服务协议缺乏定义良好的语义信息，同时也缺乏 Web 服务相互交互的表达能力，因而并不能满足 Web 服务自动发现、执行、合成、监控和恢复的需求。为了能够让 Web 服务之间互相理解各自的内容、功能以及属性，我们需要找到一种在 XML 基础之上包含语义信息的描述 Web 服务的语言。

下一代 Web 是语义 Web (Semantic Web)，其目标是为了让计算机能够明确地解释执行任务，而这正好能够解决 Web 服务自动实现的难题。实现语义 Web 的一个关键因素是要开发一种丰富的、能够很好地编码和描述 Web 内容的语言，这样的语言必须能够定义良好的语义，对描述复杂的交互和 Web 对象之间的制约关系有充分的表达能力，同时在可接受的时间和资源的限制下，能够自动操作和推理。目前已有的语义 Web 语言均基于 XML，包括 RDF(Resource Description Framework)、RDF Schema、DAML+OIL，以及最近发布的 Web Ontology Language (OWL)。DAML+OIL 和 OWL 是建立在人工智能知识表示的基础之上的语言，它们提供了一种自然的方式来描述在 Web 词典间的类与子类之间的关系，以及在类与类之间（或子类与子类之间）关系上的限制。

智能 Web 服务的目标就是以一种明确的、计算机能够理解的语言来描述 Web 服务的功能和内容，同时增强已经存在的 Web 服务操作的性能和鲁棒性，比如 Web 服务的发现和调用。智能 Web 服务也将使大量的自动化任务成为可能，包括自动合成、交互、运行监控和恢复。

目前国内外在智能 Web 服务研究方面主要着眼于两个方面：一方面是创建一种计算机之间能够互相理解的并能充分表示 Web 服务的内容、功能、属性、接口以及规则和限制条件的语言；另一方面是在这种语言基础之上提出一种使 Web 服务之间能够实现自动发现、选取、执行、合成以及交互的模型或体系结构。

现有的 Web 服务的工业标准（诸如 SOAP、WSDL、UDDI 等）能否用来描述智能 Web 服务呢？实际上这些工业标准主要能够实现 Web 服务的自动发现和调用，但是由于这些工业标准缺乏语义信息，因此无法实现 Web 服务的自动交互和合成。因此大家的研究方向都着眼于目前语义 Web 领域比较成熟的语言（如 RDF、RDF Schema 和 DAML+OIL），希望能够在此基础上开发一种新的满足智能 Web 服务需求的语言。

智能 Web 服务语言

在已经发布的智能 Web 服务语言中，DAML-S 语言得到众多研究者的认可。它是一种由 DARPA 资助的由多家研究机构共同创建的用于描述智能 Web 服务的 DAML+OIL 本体语言。

在 DAML-S 的官方网站是这样描述 DAML-S 的：DAML-S 为 Web 服务供应商提供了一套核心的标记语言集，使之可以以一种明确的、计算机能够解释执行的方式来描述 Web 服务的属性和功能。

DAML-S 是在若干 Web 服务工业标准之上开发的，同时加入了丰富的类型和类信息，我们可以使用这些信息来描述和限制 Web 服务。而且它采用一种获取 Web 服务的控制流和数据流的处理模型，集成了更多的类表示。这样的语言能够把 Web 服务聚合成分类的层次结构，并且还带有类以及类实例之间关系和限制的丰富定义。这种定义良好的语义信息通过现有的功能强大的推理技术，就能实现对这些结构的自动操作。

简而言之，DAML-S 使语义 Web 服务的自动交互成为可能。设计 DAML-S 是针对语义 Web 服务自动执行的任务，包括自动发现、自动调用、自动交互、自动集成、自动执行监控与恢复、自动模拟和验证。

DAML-S 主要由三部分组成，分别是

Service Profile: 描述 “What the services does”，指明所描述的 Web Services 的功能与接口，以便于服务代理能够搜索与匹配该 Web 服务；

Service Model: 描述 “how the services works”，指明当服务被调用时的操作，以便服务代理进一步匹配，以及服务合成和服务的协调工作和监控；

Service Grounding: 指定调用服务的具体细节，比如通信协议，调用端口等。

目前已经有一些研究机构推出了智能 Web 服务的模型或者调用流程。其中最受推崇的是由斯坦福大学 McIraith 博士提出的一种智能 Web 服务实现模型。该模型建立在 DAML-S 语言之上的，也就是说，提供 Web 服务的系统都采用 DAML-S 来描述自身的 Web 服务，各个系统之间也通过 DAML-S 进行信息交换及服务合成。

如何把工业界和学术界对 Web 服务的研究结合在一起，把智能 Web 服务的研究成果尽快地付诸于商业实践，是目前各大 Web 服务公司正在研究的课题之一。

Web 服务的实质就是“共享资源、交互通信、协作研究”。通过分布在全球各地 Web 上的各种资源在某种共同认可的协议下达到共享的目的，当然，“服务”也是一种资源，它的内涵十分广泛，从计算服务、查询与检索、信息处理都可以归结为“服务”，而智能 Web 服务就是使得这种共享更加人性化、定制化、协调性、合理性，随着智能 Web 服务技术研究的不断深入，社会的信息处理将进入一个可以自动完成 Web 服务查找和合成的智能 Web 服务时代。

尤其在构建实用的电子商务/电子政务系统中，智能 Web 服务将会起着关键作用。其实，无论什么信息系统，把“广义的服务”作为资源来考虑，在“Web 服务”的架构下，达到共享的目的，这是现今信息社会的发展趋势。

相关链接 Web 服务的定义

目前比较统一的对 Web 服务的定义是由 IBM 公司提出的：“Web 服务是独立的、模块化的应用程序，能够在网络（一般是 WWW）上被描述、发布、查找和调用。”

在 Web 服务中存在三种角色：服务提供者（Service Provider）、服务请求者（Service requester）和服务代理（Service broker）。服务提供者向服务代理发布其能够提供的服务，当服务请求者发出服务请求时，服务代理负责寻找对应的服务并提供给服务请求者，其交互过程如图所示。

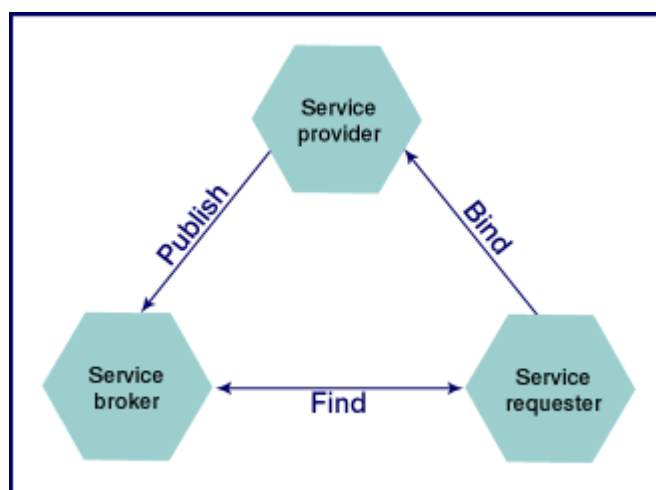


图 Web 服务调用过程