

# 基于 MAS 的智能用户界面建模

刘 扬, 郑逢斌

(河南大学计算机与信息工程学院, 开封 475004)

**摘要:** 提出一种基于多 Agent 系统(MAS)的智能用户界面分层模型和交互框架。采用基于互信息的“化学键”模型解决多通道融合和多媒体展现问题, 使用基于环境感知的“变色龙”机制实现移动的普适访问策略, 利用经济学“市场”模型和 Web 服务技术实现 P2P 协同网格计算, 并给出系统动态交互模拟的化学抽象机定义。应用实例表明, 该模型能提供自然高效的交互模式, 减少人机隔阂。

**关键词:** 智能用户界面; 多通道交互; 多 Agent 系统; 网格计算; 普适计算

## Intelligent User Interface Modeling Based on MAS

LIU Yang, ZHENG Feng-bing

(School of Computer and Information Engineering, Henan University, Kaifeng 475004)

**【Abstract】** Hierarchy model and interaction frameworks of intelligent user interface based on Multi-agent System(MAS) are proposed. The problems of multi-modal input information fusion and multimedia information presentation are solved by the chemical bond theory based on mutual information. The mobile pervasive access is disposed by the mechanism of chameleon based on environment awareness, and the cooperative and effective grid computing of peer-to-peer is implemented by using the economic market model and Web service technology. The definition of chemical abstract machine for dynamic interaction simulation of the system is given. The application example indicates that this framework can provide natural and high efficiency interaction pattern, and deduce the barriers between human and computer.

**【Key words】** intelligent user interface; multi-modal interaction; Multi-Agent System(MAS); grid computing; pervasive computing

### 1 概述

人机交互(Human-computer Interface, HCI)是21世纪信息领域需要解决的重大课题之一, 用户界面(User Interface, UI)是HCI的重要媒介。自计算机诞生初的纸带I/O起, UI的发展经历了巨大变化, 其过程如图1所示。

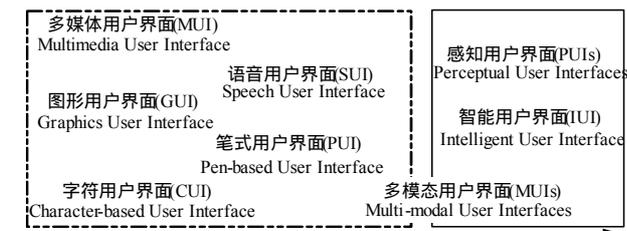


图1 UI的发展

从CUI到MUI, 从PUIs到MUIs。MUI侧重于感觉媒体的展现, 解决表示媒体到感觉媒体的转换, 包括: GUI, SUI, PUI。PUIs使HCI不再依赖于固定位置的鼠标和键盘, 模拟感知推理功能实现交互信息的理解。MUIs则侧重于多个感觉通道的输入, 旨在利用多通道的互补特性进行双向交互。智能用户界面(Intelligent User Interface, IUI)是自然交互技术取得突破之前的一种过渡, 它结合人工智能的概念、推理和表达, 按照用户、领域、任务、对话及媒体模型实现交互, 能有效地减轻用户负担。

建立符合人际交往习惯的UI是解决系统应用瓶颈的重要目标。以设备为中心的传统UI设计, 操作与内容分离, 偏离人们日常交往的习惯。复杂的UI往往以增加用户的认知负担为代价, 而简易的UI一般效用较低, 无法适应高性能网络

计算的三维虚拟交互以及移动嵌入式设备普适计算的多分辨率编码。人际交流主要依靠视觉、听觉和触觉进行, 一个理想的UI设计能模拟人际交流模式, 提供和谐自然、方便灵活、美观高效的HCI形式。为此, 提出一种基于MAS的IUI计算模型, 它采用图像、语音等多通道输入, 能普适访问高性能计算, 且易于操作。

### 2 IUI交互框架描述

鉴于交互系统的复杂性, 将该模型分为6层: (1)设备交互层; (2)媒体接入层; (3)网络架构层; (4)安全认证层; (5)智能计算层; (6)应用接口层。

该模型由多通道输入Agent(MIA)、多媒体输出Agent(MOA)、普适网络接口Agent(PCA)、UI合成Agent(ICA)、任务管理Agent(TMA)、网格计算Agent(GCA)、识别分类Agent(RCA)等Agent构成。Agent最基本的功能是从环境中感知信息, 并产生相应动作, 一个基本的Agent可用代码描述为

```

Public class AgentName {
Public PerceptInterface (); //感知与通信接口描述
KnowledgeRules(); //规则与知识描述
StateTaskSequence(); //状态与目标任务队列描述
ActionSet(); //分析处理动作集合
Private CapacitySolver();
}
  
```

**作者简介:** 刘 扬(1971 - ), 男, 讲师、硕士研究生, 主研方向: 智能信息处理, 图像与语音识别, 多媒体技术, 网络技术; 郑逢斌, 教授、博士

**收稿日期:** 2007-12-10 **E-mail:** ly.sci.art@hotmail.com

//推测、决策、控制、合作与学习等思维能力}  
IUI的分层模型以及交互信息描述分别如图2、图3所示。

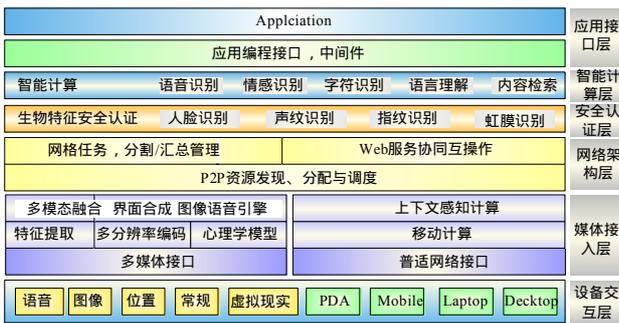


图2 IUI 分层模型

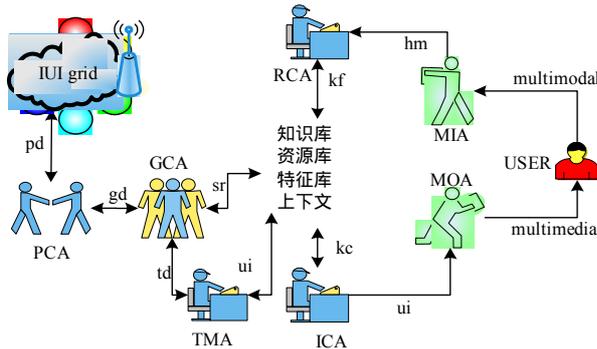


图3 IUI 交互信息描述

在图3中,各数据元素的含义分别为人机指令 hm;媒体特征知识 kf;上下文信息 kc;UI信息 ui;普适计算信息 pd;网络计算信息 gd;共享资源 sr和任务数据 td。

MIA解决了多通道特征层的多模态数据融合问题,实现智能化的自然感知功能。MOA负责虚拟现实三维音视频多媒体的展现以及便携嵌入式普适设备I/O的多分辨率编码功能。若定义用户交互的感觉媒体为 $A_i, V_i, T_i$ ,上下文定义为 $C_i$ ,则MIA在RCA的配合下,从输入通道交互信息中提取用户指令 $CMD_i$ 的感知,即:

$$CMD_i = Dec(\{A_i, V_i, T_i, C_i\})$$

而MOA在ICA的协助下,解决了用户指令反应 $ACT_i$ 的多媒体信息合成问题,即:

$$\{A_i, V_i, T_i\} = Syn(ACT_i, C_i)$$

借用“化学键”(chemical bond)理论,模拟离子键和共价键的形成机制,将系统的MIA多通道用户指令感知和MOA多媒体信息合成分别模拟为化合物的化合和分解。这里,离子键和共价键可用媒体间的互信息 $MI$ (Mutual Information)来描述。设数据信息熵 $H$ ,则媒体 $A$ 和媒体 $V$ 之间的互信息为 $MI(A, V) = H(A) + H(V) - H(A, V)$

普适网络接口Agent(PCA)是一种反应型Agent,它实现网络的接入。普适计算改变了传统桌面HCI的模式<sup>[1]</sup>,可整合无线网、有线网的功能和资源,为用户提供一致性的、不间断的、自适应的移动访问环境。PCA自动感知接入环境的变化,这里,模拟“变色龙”(chameleon)机制,以不变的基本功能组合来适应环境变化,同时,实现移动环境的任务迁移、位置感知以及上下文感知的自适应。若定义PCA所具备的各种可自由组合功能为 $F = \{D_1, D_2, \dots, D_n\}$ ,则节点所能适应的环境 $E_i$ 为 $E_i = \{PP_i, CI_i, PD_i, NP, TM, NT, TS\}$ ,其各个元素分别表示节点位置、上下文、节点设备类型、网络协议、网络传输媒体、网络拓扑结构和时态信息。PCA对环境 $E_i$ 的反应 $CA$ 可形式化

描述为

$$CA = FitF(F|E_i)$$

ICA是UI合成器,依据认知心理学和美学模型,实现展现媒体的界面合成。它由语音引擎和图像引擎根据配色集与构图模板实现任务UI的生成。根据语音和人脸情感特征,情感识别Agent(ERA)识别出用户的情感状态,使合成的UI模板能与用户的情绪相适应。UI可用四元组 $\langle TI, UI, Em, Op \rangle$ 描述,其元素分别为任务ID、UI资源ID、情感状态和任务对应操作。

TMA在网络架构层根据自身情况和网络性能来实现任务管理,具体有任务的分割、合成、调度,以及资源分配和负载均衡。网络计算任务和资源分配策略一般采用仿生群体智能协同计算技术,目的是使整体效益最大化,协同完成同一任务。MAS协作过程也可基于经济学的市场模型<sup>[2]</sup>,市场由商品、消费者、生产者和市场规划等因素组成,根据市场机制和拍卖规则,市场的最优状态是竞争平衡,即所有Agent都满意的状态为最优配置。设单节点提供的最大服务性能为 $g_i$ ,实际提供的服务能力为 $G_i$ ,实际消耗的服务量为 $S_i$ ,则整个网格可提供的计算能力 $G = \Sigma(G_i) - \Sigma(S_i)$ 。定义当前网格服务价格 $P = G / \Sigma(g_i)$ ,若 $P'$ 为上次请求时的服务价格,对于每个节点就是要在网格最大服务能力 $\Sigma(g_i)$ 不变的情况下,根据自己剩余计算能力 $g_i^*$ 去动态调整 $S_i$ 和 $G_i$ ,最终达到网格上的服务提供量和消耗量的动态平衡。其中, $g_i^* = (g_i + (S_i - G_i) / P') / P$ 。

GCA在TMA和PCA的协助下,负责网络通信和协同计算,实现资源发现与调度、节点管理及网络存储等任务,从而达到高性能计算的目的。网络资源管理要依据系统规划自寻优,当个别节点出现故障时,能及时诊断,必要时利用任务冗余功能来实现异常的恢复。如图4所示,网格节点由计算型节点和非计算型节点组成一种混合的P2P结构,P2P可将网络中不同的计算机连接起来,充分利用和共享网络中任何闲置的资源<sup>[3]</sup>。

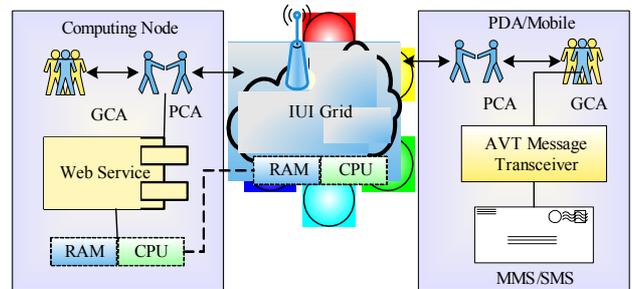


图4 IUI 网络结构

计算型节点在接收服务的同时,还要求对外提供服务;非计算型节点可在移动环境中,用类似短消息(SMS)或多媒体消息(MMS)与网格通信。系统存在一个索引服务器,负责记录并维护参与计算的所有节点信息,即所谓的公告“白板”。节点生存状态采用“心跳模型”动态进行,并在“白板”上登记,但MAS之间通信时,并不需要通过索引服务器进行,即索引服务器只起联络作用。计算节点采用Web服务来封装本节点可用资源,对外提供计算和存储功能<sup>[4]</sup>。MAS协同操作网格上的存储和计算资源,通过各个节点来调用彼此的Web服务,实施网格计算。节点的资源分配由节点的性能动态调整,设当前共有 $N$ 个节点参与,节点可提供的服务性能参数为节点的主频 $F_i$ ,可用内存 $M_i$ ,CPU可用率 $C_i$ 和网络服务QoS情况 $Q_i, i = 1, 2, \dots, N, a_1, a_2, \dots, a_4$ 为权重,则第 $i$ 个节点的

归一化性能 $f_i$ 定义为

$$f_i = (a_1 F_i + a_2 M_i + a_3 Q_i + a_4 C_i) / \sum (a_1 F_i + a_2 M_i + a_3 Q_i + a_4 C_i)$$

工作在智能计算层的 RCA 是具备学习和推理能力的识别分类器。鉴于识别分类的复杂性, RCA 既要功能分区、各自独立, 又要彼此联系, 并负责非结构化媒体对象的特征提取与模式识别, 具体包括 OCR, ASR, NLP 及视听语义特征的提取。RCA 学习到的有关 MDF, CBR, 以及图像与语言理解等知识规则, 由知识库和特征库存储。此外 PCA, MIA 和 MOA 属于静态 Agent, 一般驻留在终端节点上。其余几种 Agent 都属于移动 Agent, 当节点性能不足或有故障时, 可转移到网格中的其他节点继续进行计算。

为了描述系统交互过程, 这里采用化学抽象机 (Chemical Abstract Machine, CHAM)<sup>[5]</sup> 模型来动态模拟交互状态。设节点内部 Agent 的通信采用消息 ( $mr, mw$ )、管道 ( $pr, pw$ ) 和基于共享内存的“黑板”通信 ( $br, bw$ ) 3 种机制。首先定义分子 M、处理元素 P、数据元素 D、连接元素 C 的语法结构:

$$M ::= P|C \diamond M|M \diamond C|M \diamond M$$

$$P ::= MIA|MIA|RCA|TMA|ICA|GCA|PCA|DB|IUI\_Grid$$

$$D ::= hm|kf|kc|ui|pd|gd|sr|mc|td$$

$$C ::= mr(D)|mw(D)|br(D)|bw(D)|ir(D)|or(D)|pr(D)|pw(D)$$

中缀符  $\diamond$  表示处理元素 I/O 行为的状态, 其余数据元素含义分别为人机指令  $hm$ 、媒体特征知识  $kf$ 、上下文信息  $kc$ 、UI 信息  $ui$ 、普适计算信息  $pd$ 、网格计算信息  $gd$ 、共享资源  $sr$ 、管理控制信息  $mc$  和任务数据  $td$ 。然后定义初始溶液  $S_0$  和反应规则  $T_i (i=1, 2, \dots)$ :

$$S_0 = \{MIA \diamond mw(hm), pr(ui) \diamond MOA, mr(hm) \diamond br(kf) \diamond bw(kf) \diamond RCA, br(kc) \diamond pw(ui) \diamond bw(kc) \diamond ICA, br(mc) \diamond mr(td) \diamond bw(mc) \diamond mw(td) \diamond TMA, mr(pd) \diamond mr(gd) \diamond mw(pd) \diamond mw(gd) \diamond PCA, mr(pd) \diamond mw(pd) \diamond IUI\_Grid, br(sr) \diamond mr(gd) \diamond mr(td) \diamond bw(sr) \diamond mw(gd) \diamond mw(td) \diamond GCA, ir(kf) \diamond DB, DB \diamond or(kf), ir(sr) \diamond DB, DB \diamond or(sr), ir(mc) \diamond DB, DB \diamond or(mc), ir(kc) \diamond DB, DB \diamond or(kc)\}$$

$$T_1 = m \diamond o(d) \rightarrow o(d) \diamond m$$

$$T_2 = mr(d) \diamond m_1, mw(d) \diamond m_2 \rightarrow m_1 \diamond mr(d), m_2 \diamond mw(d)$$

$$T_3 = pr(d) \diamond m_1, pw(d) \diamond m_2 \rightarrow m_1 \diamond pr(d), m_2 \diamond pw(d)$$

$$T_4 = br(d) \diamond m_1, bw(d) \diamond m_2 \rightarrow m_1 \diamond br(d), m_2 \diamond bw(d)$$

$$T_5 = br(d) \diamond m, or(d) \diamond DB \rightarrow m \diamond br(d), DB \diamond or(d)$$

$$T_6 = bw(d) \diamond m, ir(d) \diamond DB \rightarrow m \diamond bw(d), DB \diamond ir(d)$$

$$T_7 = MIA \diamond mw(hm), mr(hm) \diamond br(kf) \diamond bw(kf) \diamond RCA, DB \diamond or(kf) \rightarrow mw(hm) \diamond MIA, br(kf) \diamond bw(kf) \diamond RCA \diamond mr(hm), DB \diamond or(kf)$$

(上接第 232 页)

本系统应用程序采用 QT3.3.5 设计<sup>[5]</sup>, 并通过 Tslib 函数库来完成触摸屏的校正<sup>[6]</sup>。应用程序支持多线程, 主要功能包括: 图像实时滚屏显示, 图像数据存储, 向红外相机发送控制指令。各种功能通过触摸屏启动。

## 6 结束语

通过嵌入式技术构建的便携式红外相机小巧灵活, 不需要电脑的辅助, 给实验带来便捷。另外, 随着红外焦平面技术的普及, 红外相机将进一步走入民用市场, 嵌入式技术将是必然的选择。此外, 尽管该系统针对一款 1024 元的红外阵列探测器, 但其关键技术适用于其他类型的相机, 也适用于其他的嵌入式数据采集设备。其他设计需要考虑系统处理速度、数据存储速度、显示的要求等。基于这些考虑, 可以选择性能更好的微处理器和 LCD。

$$T_8 = pr(ui) \diamond MOA, br(kc) \diamond pw(ui) \diamond bw(kc) \diamond ICA, DB \diamond ir(kc) \rightarrow MOA \diamond pr(ui), br(kc) \diamond ICA \diamond pw(ui) \diamond bw(kc), ir(kc) \diamond DB$$

这里仅给出通用反应规则 (T1~T6) 和部分特殊反应规则 (T7~T8), 利用这些规则可动态模拟用户交互及 MAS 协同过程, 限于篇幅原因, 详细推演过程不再赘述。

## 3 应用实例

为验证本模型的可用性, 特开发面向儿童的实验性系统。选 8 名 3~4 岁的正常儿童进行测试, 结果表明, 儿童并不需要太多的学习即能很快掌握系统交互操作。多媒体提示能有效引导和纠正操作, 儿童对指点触摸和受限语音控制方式较感兴趣。但由于儿童操作不规范, 因此身份认证和情感识别反应不及时, 此外, 语音指令易受环境音的干扰。

## 4 结束语

该模型利用普适访问和基于 MAS 和 P2P 的网格技术, 在移动环境中, 结合 Web 服务技术, 实现资源分布共享和任务调度的无缝衔接。该模型融合多通道输入以及多媒体展现的智能交互模式, 有助于提高交互的自然性和易用性。本文提出的交互框架具有较强的通用性, 可为高效、自然的交互设计提供参考。下一步将结合一些成熟的研究成果, 继续研究基于 IUI 的具体应用平台。

### 参考文献

- [1] 岳玮宁, 董士海, 王悦. 普适计算的人机交互框架研究[J]. 计算机学报, 2004, 27(12): 1657-1664.
- [2] Sander M B, Enrico G. Market-based Recommendation: Agents That Compete for Consumer Attention[J]. ACM Transactions on Internet Technology, 2004, 4(4): 420-448.
- [3] 李飞, 何炎祥, 张晶晶. 基于性能预测和 Agent 的网格负载均衡的实现[J]. 计算机工程, 2005, 31(15): 104-106.
- [4] 易明, 金海. 基于 WSRF 的 Web 服务资源的设计[J]. 计算机工程, 2006, 32(23): 262-264.
- [5] Inverardi P, Wolf A L. Formal Specification and Analysis of Software Architectures Using the Chemical Abstract Machine Model[J]. IEEE Transactions on Software Engineering, 1995, 21(4): 373-386.

### 参考文献

- [1] Yaghmour K. 构建嵌入式 Linux 系统[M]. O'Reilly, 译. 北京: 中国电力出版社, 2004.
- [2] 王岳环, 路自谦. Linux 下实时图像传输卡驱动程序的研究与实现[J]. 计算机工程, 2006, 32(15): 271-273.
- [3] Corbet J, Rubini A, Greg K H. Linux 设备驱动程序[M]. 3 版. 魏永明, 耿岳, 钟书毅, 译. 北京: 中国电力出版社, 2006.
- [4] 陈雁飞, 赵岳松, 陈榕. 基于构件技术的嵌入式 GUI 系统[J]. 计算机工程与设计, 2006, 27(4): 561-564.
- [5] Blanchette G, Summerfield M. C++ GUI QT3 编程[M]. 齐亮, 译. 北京: 北京航空航天大学出版社, 2006.
- [6] 程龙, 刘彦明, 鲁刚, 等. Qt/E 中文环境的搭建及对常用输入设备的支持[J]. 计算机工程, 2007, 33(15): 275-276.