

智能 Agent 及多 Agent 在虚拟环境中的应用研究进展

——《人工智能》读书报告

董子龙 10421038

浙江大学 CAD&CG 实验室

摘要：

Agent 在虚拟环境中的应用研究主要是单一的对话表情功能，虚拟生物的交互运动和社会系统的内部结构。本文概括性地介绍智能 Agent 及多 Agent 在虚拟环境中的应用现状，根据系统的规模，由易到繁地说明当前研究的体系框架和实现。

关键字：智能 Agent，多 Agent，虚拟环境，体系框架

一 引言 Introduction

随着计算机硬件的不断升级，用户对系统环境的要求越高。在系统环境上，用户希望在一个与现实世界相仿甚至难辨真假的平台上工作；在交互控制上，用户希望系统提供更强大更全面的辅助，适应用户的操作。前者促进了虚拟环境技术的不断分化深入；后者常常引入人工智能领域的重要概念——Agent，在面向过程，面向实体，面向对象编程的基础上发展出更抽象的面向 Agent 编程[5]。

利用 Agent 技术建构的智能化实体，广泛应用在探测、材料、医疗、教育、游戏等领域，帮助或代替人类完成工作[7]。

1.1 一些工作 Some Works

虚拟环境涵义很宽，对现实世界中物理的方法、行为、地理和抽象的思维、情绪、感情的模拟都属于虚拟现实的范畴，所以 Agent 与之的结合点很多，在实际应用给开发人员很好的实现灵感。

*Helmut*设计一个卡通形象的动画(Animated)Agent[2]，具有让人信服的表情和社交能力，用于日本学生的英语对话训练。他们使用XML风格的MPML 脚本语言(Multi-modal presentation Markup Language)控制不参与交互的Agent角色的行为语言。*Baldi*[3]由俄勒冈州研究院、卡内基梅隆大学等联合开发的语言教学系统，是一个具有听觉和可视化演讲能力的Agent，结合了语音识别，面部动画，表情跟踪和语音朗读四个方向的技术。

*Marche*和*Anton*的*Jacob*[4]项目在虚拟现实系统中构建一个叫做*Jacob*的拟人(Human-like)Agent，传达给用户操作指示。*Jacob*项目涉及虚拟现实模型的软件工程、自然语言等交互模块和融合Agent技术，由于任务模块和指令模块的分离，*Jacob* Agent可以容易地移植到其它虚拟环境中。*Ipke*和*Yong*在*VIENA*[6]项目中(Virtual Environments and Agents)使用适应性Agent帮助用户设计和探索3D图形，这是一个多Agent系统，Agent之间根据当前状态互相协作实现用户的指令；而对于用户来说，只是面对一个能理解和执行语言指令的Agency。

Agent的研究还包括控制脚本[8][9][10][18]和实际应用的性能分析[11][12][13]。

1.2 本文框架 Layout of This Article

本文从三个层次总结智能Agent和多Agent系统在虚拟环境中的应用。第3章介绍会话(Conversational)Agent，第4章介绍行为(Behavioral)Agent，第四章关于多Agent。为了帮助读者形成完整的概念，第2章会对智能Agent做出探索性定义。

二 关于 Agent Concepts of Agent

智能 Agent 的定义很不统一，在不同应用，从不同的角度，研究人员总是在自己的理解基础上给出说明。Stan 和 Art 考察大量 Agent 的概念，得出模糊却内涵丰富的结论[14]：

自主(Autonomous)Agent系统是环境的一部分，只在特定的(Situated)环境中感知环境并作用于环境，从而实现自身的日程，影响将来的感知。自主Agent和智能Agent应该是等价的。

智能Agent总是与之对应的环境紧密结合，离开了所处的环境，Agent就失去其存在的意义，再也不是Agent。她通过感应器(Sensor)感知环境，利用效应器(Effector)作用于环境。她具有自己的信念，即有目的，有意图地行为，使环境将来的状态符合日程。

智能 Agent 可以是硬件，如机器人，人眼等，自然更是软件(Softbot)。本文主要讨论软件(Software)Agent 的设计框架和实现。

对智能 Agent 概念的详细解释可以从 Winikoff[32]等提出的 SAC(Simplified Agent Concepts)——“一个简化的模型，能让更多人开发智能 Agent 系统，但是保存 BDI 的能力和效率”——中获得。

2.1 描述性定义 Descriptive Definition

描述性定义指出只要一个基于硬件或软件的系统具有某些描述性特性，就可以被称为Agent。可分为弱定义和强定义[15]。

弱定义包括自主性(Autonomy)，社会性(Social ability)，反应性(Reactivity)等人类特有的性质。

强定义在弱定义的基础上加入知识(Knowledge)、信念(Belief)、意图(Intention)、责任(Obligation)等精神概念，有研究者称之情感(Emotional) Agent[17]。

其它很多词汇也常被研究者用来描述Agent，如移动性(Mobility)，诚实(Veracity)，善良(Benevolence)，理性(Rationality)，长寿性 (Longevity)，前瞻性(Pro-active, Goal-directed) 等。

2.2 形式化定义 Formal Definition

本文的形式化定义只是以理想理性(Ideal Rational)Agent[16]为例，抛砖引玉，帮助读者对不同类型的 Agent 形成自己的形式化方法。

理想理性 Agent:

对每一个可能的感知序列(Percept sequence)，她都能基于感知序列和内建的知识提供的证据采取行为，期望性能最优化。Agent 由 PAGE 描述：感知(Percept)，行为(Action)，目标(Goal)，环境(Environment)。

Agent 等于体系构架(Architecture)和程序(Program)。Stuart 和 Peter 提出最基本的骨架 Agent 程序：

function SKELETON-AGENT (*percept*) **returns** action

static: *memory*, the agent's memory of the world

memory ← UPDATE-MEMORY(*memory*,*percept*)

action ← CHOOSE-BEST-ACTION(*memory*)

memory ← UPDATE-MEMORY(*memory*, *action*)

return *action*

而根据程序设计的立足点不同，由简到繁区分成：表格驱动(Table-driven)Agent，简单反射(Simple reflex)Agent，跟踪内部状态的反射 Agent (Reflex agent with internal state)，基于目标的(Goal-based)Agent，基于效能的(Utility-based)Agent。

三 会话 Agent Conversational Agent

会话Agent主要在教育培训程序中扮演导师和同学的角色，随时随地供给交流学习的伙伴，增强娱乐性和参与性，极大提高教学效率，也可以在场馆中推广产品，或说明旅游项目。Microsoft Office系列软件的帮助精灵是我们最熟悉的会话Agent。会话Agent具有一般的语言能力，也有一定的声音识别能力，甚至有感情；在虚拟环境中，她常常通过图形具体化(Embodied)。本章将介绍Baldi和Max两个系统，说明如何让Agent表示出表情和嘴形，或手势，进一步加入情感。

关于会话Agent的更多研究请参考[22][23][24][25]。

3.1 表情和嘴形 Facial Expression and Lips

与一个表情丰富、语言流畅的伙伴交流，无疑是长期在计算机前学习生活的人很吸引的事。

图 3-1 所示的 Baldi[3][19]曾在上文引述过，是一个由计算机驱动的聊天人头，暂时用于聋哑儿童的课堂语言教学。她的存在和功能完全依赖于计算机动画控制，和文本—语言合成。

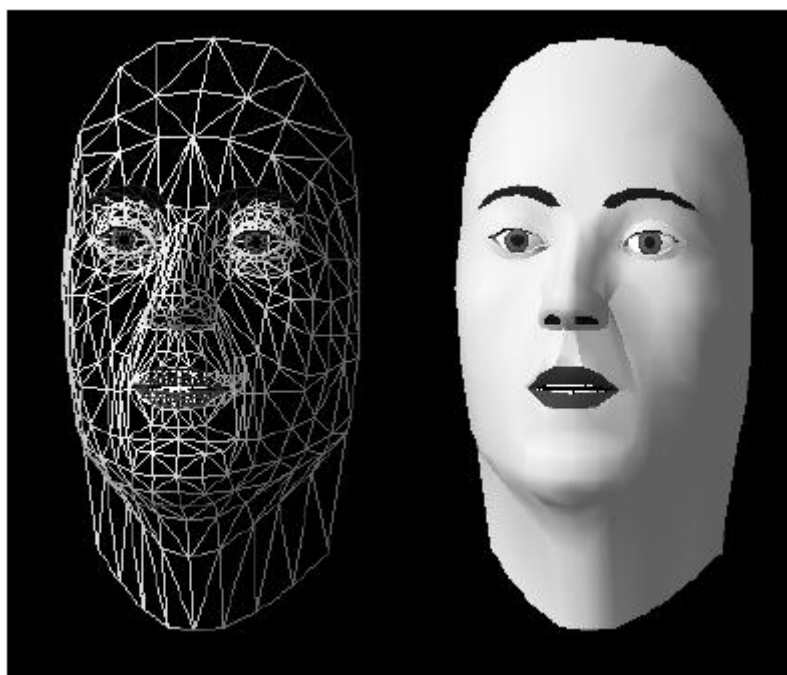


图 3-1

她的语言有33个参数：颌旋转和挤压，嘴的水平宽度，嘴唇弯曲和突出控制，下嘴唇褶皱，嘴唇垂直位置，牙齿的位移，舌头的角度、宽度和长度。

为了能在低端设备上实时绘制，研究者采用目标相似合成(Terminal Analogue

Synthesis)技术，仅令最终结果看起来很像，并没有完全模仿生理结构。大约900个多边形边边相连组成Baldi的眼睛、瞳孔、虹膜、巩膜、眉毛、鼻子、皮肤、嘴唇、舌头、牙齿、脖子。多边形的拓扑结构和动画由一个参数集控制。

Baldi的研发进展主要在控制参数的增加和修改，舌头的两代实现方式，视觉语言合成控制，文本—语言合成，双模（视觉/听觉）合成，和并行信息处理。最终大概有20000行C代码，可在SGI和PC上实时运行，图3-2是一些结果。

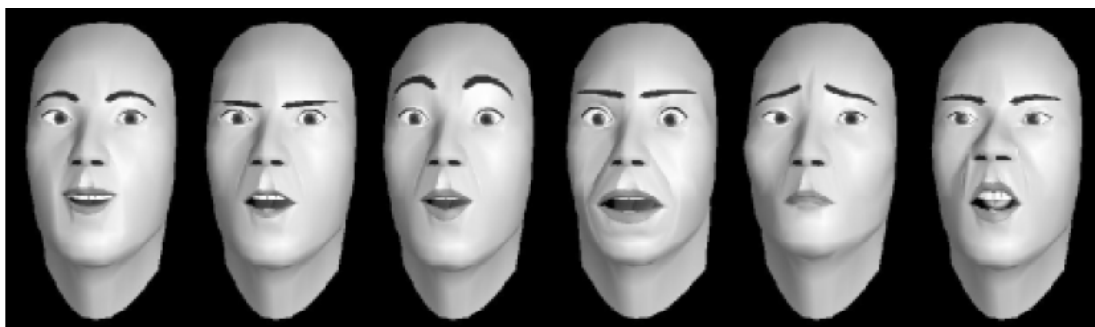


图3-2 Baldi的表情：高兴、生气、吃惊、恐惧、伤心、恶心

3.2 手势 Gesture

手势是人们自发的无意识的语言表达，是很重要的沟通手段，对多模式 (Multi-Modal)会话 Agent 的具体化最终都需要一套肢体语言的支持，可从真人捕捉，或手动预定义。

拟人 Agent Max[20]是在 3D 虚拟环境中的装配专家，通过上肢活动和口述的相互协同向用户展示装配过程，图 3-3。

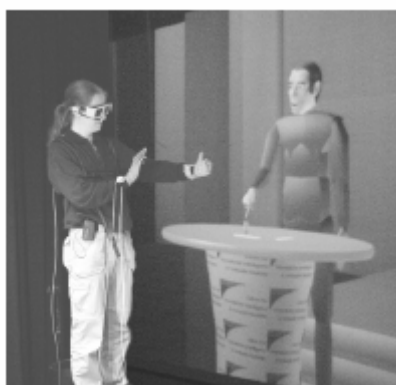


图 3-3 与 Max 多模交互

Max的多模(Multi-Modal)发言直接由一种基于XML的语言描述，包括语言词句和非语言行为。语句被时间点(Time Point)分隔，特定的手势动作被定义在相应的时间间隔中。手势由发出阶段(Stroke Phrase)的时空特性决定，由子动作组成。

子动作被定义成：(1)拳头的位置，(2)手的形状，(3)拳头的朝向；每一项性质用数值或符号表示。子动作是静态或动态，静态代表一定时间内不变，动态子动作又由连续的动作片断(Segment)组成。手势的特点，如并发(Simultaneity)，滞后(Posteriority)，重复(Repetition)和对称(Symmetry)，既可以显示说明，也可以用行为通信函数表达。图3-4是XML说明片断，图3-5是其结果。

```
<definiton><utterance>
  <specification>
    And now take <time id="t1"/> this bar <time id="t2" chunkborder=
      "true"/> and make it <time id="t3"/> this big.<time id="t4"/>
  </specification>
  <behaviorspec id="gesture.1"> <gesture>
    <affiliate onset="t1" end="t2"/>
    <function name="refer_to_loc">
      <param name="refloc" value="$Loc-Bar.1"/>
    </function>
  </gesture> </behaviorspec>
  <behaviorspec id="gesture.2"> <gesture>
    <affiliate onset="t3" end="t4"/>
    <constraints>
      <symmetrical dominant="rightArm" symmetry="SymMS">
        <parallel>
          <static slot="HandShape" value="BSflat(FBround all o)
            (ThCpart o)"/>
          <static slot="PalmOrientation" value="PalmL"
            mode="absolute"/>
          <static slot="ExtFingerOrientation" value="DirA"/>
          <dynamic slot="HandLocation">
            <segment type="linear">
              <value type="start" name="LocShoulder
                LocCenterRight LocNorm"/>
              <value type="direction" name="DirR"/>
              <value type="distance" name="125.0"/>
            </segment>
          </dynamic>
        </parallel> </symmetrical> </constraints>
      </gesture> </behaviorspec>
    </utterance></definiton>
```

图 3-4



图 3-5

基于特点的手势动画要求运动规划和上肢控制。在高级(Higher-Level)规划时，手势发出(Stroke)阶段的约束条件被充分限定，按时序传给独立的运动控制模块(Motion Control Module)，手、拳头、手臂、脖子和脸都有不同的模块。低级(Lower-Level)规划采用局部运动程序(Local Motion Program)，几个 LMP 组成

的运动程序根据动力学原理控制上肢的移动。**LMP** 定义在外部坐标系或关节连接角上，根据当前运动状态自动激活，并按照给定的边界条件互相连接，如图 3-5。

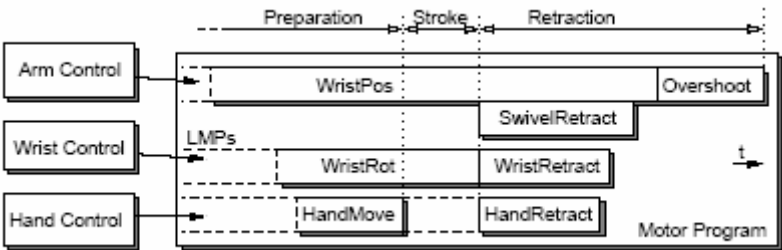


图 3-5

在介绍口语和手势同步协作之前，先说明块(Chunk)的意义。手势被分成准备(Preparation)、保持(Hold)、发出(Stroke)、收回(Reaction)四个阶段，语言也组织成音调阶段(Intonational Phrase)。块就是一个音调阶段加上一个手势阶段，复杂的发言和手势由多个块组成。

在 Max 的同步结构中，块各自在独立的黑板上交叉执行。在 InPrep 状态，口语合成和手势规划模块共同决定完成规划(Planning)过程；规划完成后，块进入 Pending 状态，如果上一块在 Subsiding 状态，调度器将切换到处在 Lurking 状态的块，接着进入 InExec 状态，最后，如果还有 LMP 活动，到 Subsiding 状态，否则 Done，图 3-6。

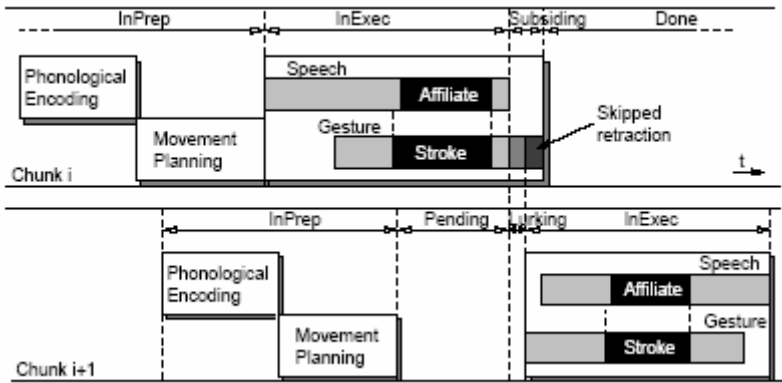


图 3-6

3.3 情感 Emotion

情感为人类特有，在 Agent 的应用中，如何给 Agent 赋予一定的感情

(Emotion)、情绪(Mood)和个性(Personality)一直是研究者们最感兴趣的问题。情感 Agent 更有生命，更让人信服，Max[21]在这个方向上又前进了一步。

为了表现情感，需要：

- (1) 评估内外事件对动态情感(Emotion Dynamics)的影响；
- (2) 模拟 Agent 的感情和情绪，及其相互作用；
- (3) 确信将 Agent 的情感表达出来。

第一个方面可以：在非认知(Non-Cognitive)层，感应器从虚拟现实系统直接反馈；或在认知层，BDI 理解器的意愿推理得到。下面详细说明另外两个方面。

3.3.1 情感系统的内部结构 Internal Structure of Emotion System

首先，区分感情(Emotion)和情绪(Mood)两个概念。感情由激励引起，持续时间短，情绪则比较均匀的分布在时间维度上，时间比较长；感情能加强或减弱情绪，而情绪可看作比较弱而且较长久的感情。

Max的感情和情绪分布在二维空间上(图 3-7)，感情价(Valence)对应X轴，情绪价对应Y轴。原点被认为是稳定点，所以类比弹簧，感情价和情绪价分别对应一个弹性系数 d_x 、 d_y ，由此产生的力 F_x 、 F_y 使它们在一定时间后恢复到原点。

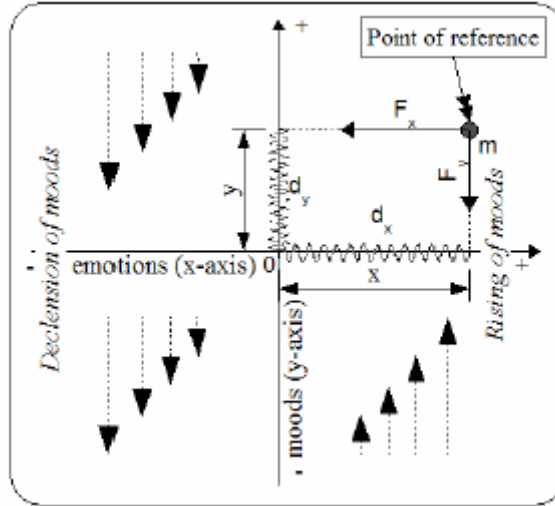


图 3-7

在公式 1 中，将感情当作情绪的变化梯度，可以表达感情对情绪的增强和减弱作用。较小的 a 导致惰性，较大的 a 使 Agent 更加情绪化。

$$\frac{\Delta y}{\Delta t} = a \cdot x$$

公式 3-1

作为对动态情绪的补充，厌烦(Boredom)被添加到系统中。如果Agent的情感处在原点的一个椭圆区域(图3-8)内，在没有刺激的情况下，厌烦程度线形增长；一旦出了这个区域，厌烦程度重设为0。厌烦线形增长公式是： $Z(t+1) = Z(t) - b$ ，Z定义在 $[-1,0]$ ，即-1时，Agent最感厌烦。

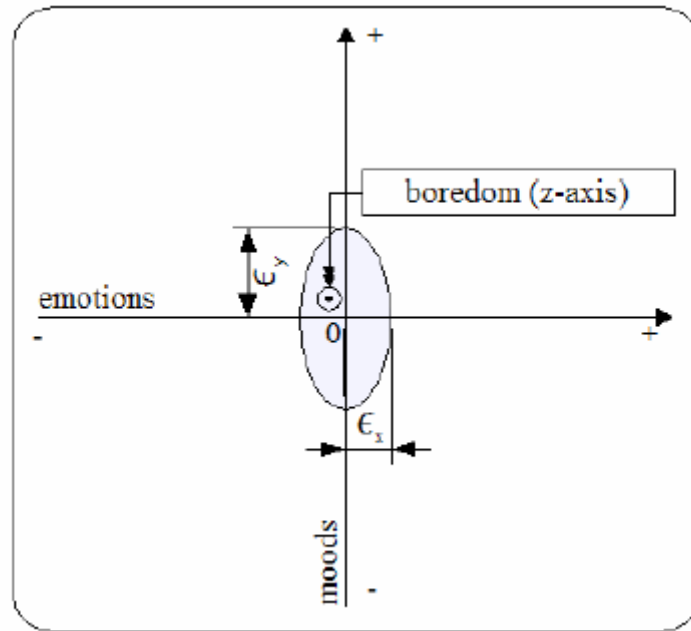


图 3-8

以上的模型有效的概括了人类的情感，然而体现在交流中，情感有三个维度内涵：愉快(Please)、激活(Arousal)、支配(Dominance)。感情、情绪、厌烦最终被映射到 PAD 空间，并归类。公式 3-2 规定映射方式，定义域由公式 3-3 确定。 x_t 、 y_t 、 z_t 分别代表感情价、情绪价、厌烦程度， t 是时间。

$$K(x_t, y_t, z_t, t) = (p(x_t, y_t), a(x_t, y_t), d(t)),$$

$$\text{with } p(x_t, y_t) = \frac{1}{2} \cdot (x_t + y_t) \quad \text{and} \quad a(x_t, z_t) = |x_t| + z_t$$

公式 3-2

$$D(t) = (x_t, y_t, z_t), \quad \text{with } x_t = [-1, 1], y_t = [-1, 1], z_t = [-1, 0]$$

公式 3-3

$p(x_t, y_t)$ 是愉快度。愉快全面代表价信息，当感情价和情绪价都是最大值时，Agent最愉快，反之则最勉强。

$A(x_t, z_t)$ 是激活度。激活度总是处在困倦，及心理敏感度和生理兴奋度最

大值之间。总是取负值的厌烦程度起降低激活度的作用。

$d(t)$ 是支配度。支配的存在是为了区分影响外界状况和事件的感觉，如厌烦和生气，或被外界环境控制的感觉，如恐惧和悲伤。

根据以上的计算结果，将情感归入相应的类，如图 3-9。一个形容词代表一种 PAD 类别，人工设定相应的 PDA 值。

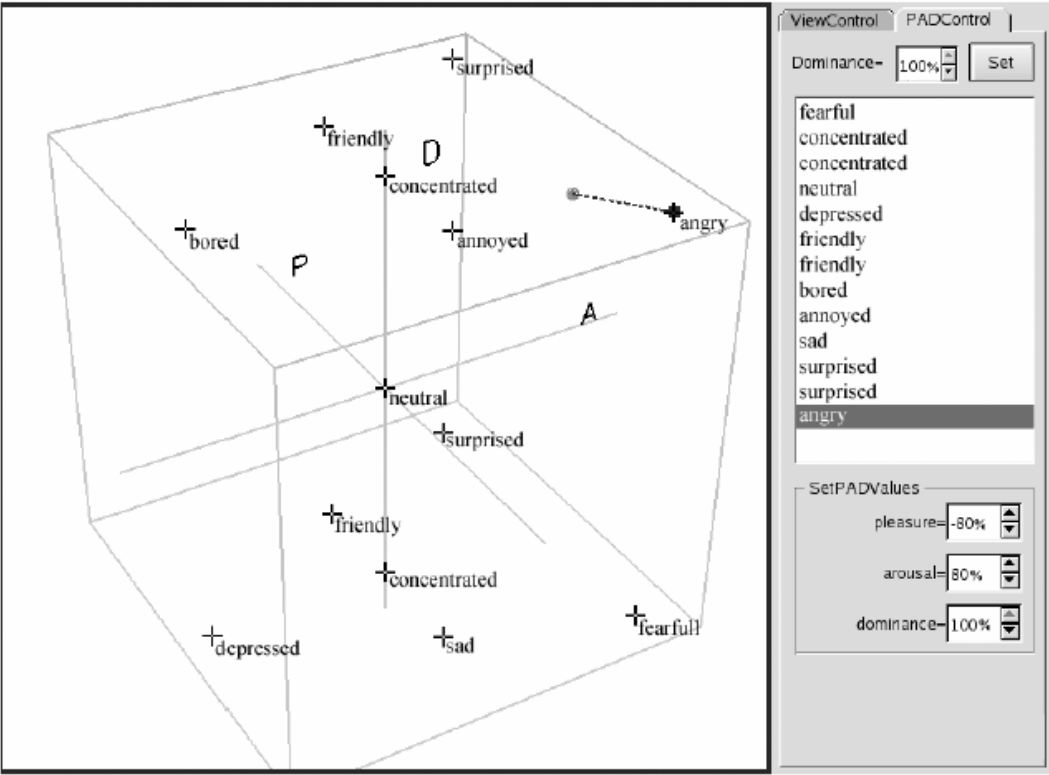


图 3-9

每一类对应一个激活阈值 Φ 和一个饱和阈值 Δ ，如图 3-10。当点在且仅在一个类的 Φ 内，类被激活，激活程度由公式 3-4 计算。当点在且仅在一个类的 Δ ，类完全激活， $W = 1$ 。如果以上条件都不满足，Agent 进入特殊的混乱 (Confused) 状态。

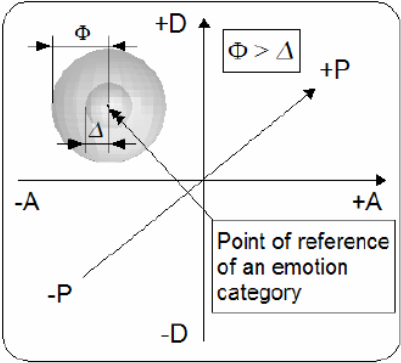


图 3-10

$$w = 1 - \frac{d - \Delta}{\Phi - \Delta}$$

公式 3-4

3.3.2 集成和应用 Integration and Application

最终，将情感系统集成到现有的 Max 中，设定相应的输入输出，图 3-11。

Max 的认知构架通过两种途径提供给情感系统情感价信息：

- (1) Interpretation 和 Dialog Manager，审察过程；
- (2) Perception，对正面或负面刺激的直接反应。

而情感系统反馈给认知构架三方面数据：

- (1) 情绪价和厌烦程度；
- (2) PAD 三元组；
- (3) 被激活的情感类型及其强度，或者混乱(Confused)。

前两项数据用于 Behavior Generation，第三项用于 Dialog Manager。

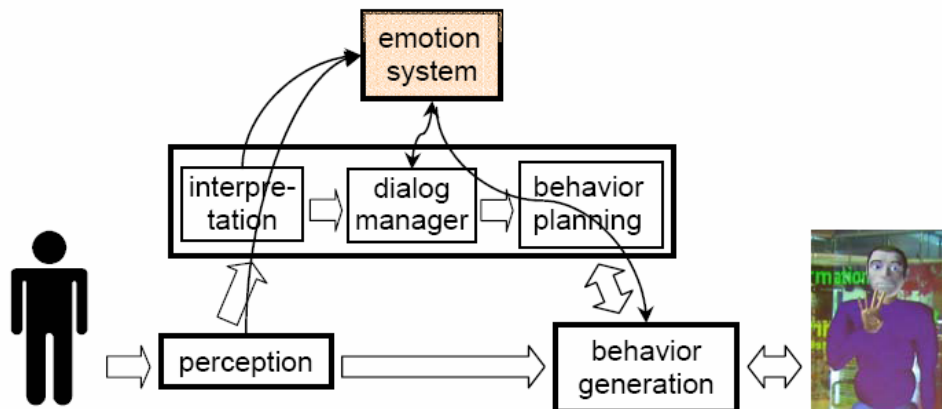


图 3-12

图 3-13 是 Max 生气并走出画面的例子。



图 3-13

3.4 小结

会话 Agent 对实时性的要求比较高，在交互、展示、教育上有广范的应用，但是由于硬件的限制，Agent 的外形一般计较粗糙。

另一方面，会话 Agent 的语言、行为、情感都是由开发者根据实际环境预先设定，缺少自学习的能力，不能一般地在各种情况中切换。

还有，语言识别模块是会话 Agent 比较薄弱的一块，这个模块如果不能很好的适应环境，会使 Agent 答非所问，完全背离设计的目的。

所以，将来会话 Agent 的发展应该至少有三个方向：

- (1) 更逼真的视觉效果，包括精细的肌肉变化、眼神的变化、毛发的飘动等，也许可以用骨架来控制肌肉变化；
- (2) 学习能力的增强，对新语句的融会和理解，这需要语言学和人工智能其它领域的支持；
- (3) 语言识别模块更强大，声音的采集处理，能够应付各种各样的意外因素。

四 行为 Agent Behavioral Agent

行为 Agent 与会话 Agent 有所不同。行为 Agent 研究的是 Agent 在虚拟环境中根据环境的状态，调整行为，规划运动，完成目的，主要用于控制动物在虚拟环境中的运动。

徐晓媛[26]和James[27]分别就动物和人的运动模拟所做的博士工作，深入细致的探索了理论模型和实践方法。[28]提出垂直的多级抽象层次结构AHA (Asynchronous, Hierarchical Agents)，模拟人类多层次的迭代思考过程。[29]将增

强学习法(Reinforcement Learning)用于自发虚拟Agent(Autonomous Virtual Agent)的运动控制，采用Q-Learning和TD-Learning两种算法。[30]的离散调度机制决定运动随时间的变化，而由一个多通道构架完成动作的执行和渲染。

MIT Media Lab C4[31]的理念基于大脑被分成很多系统，并在黑板上交流。主要模块有：感应系统和感知系统，动作系统，导航和运动系统，短期记忆和心理黑板(Mental Blackboard)。C4 的工作集中在训练一只家庭宠物上，4.1 详细介绍。

4.2 介绍 Wen[32]等提出的将 Agent 的行为可视化的方法，。

4.1C4

C4 描述了一种新的多层次脑体系，模拟生活在图形世界中的自主半自主生物。设计目标有：反映行为，学习能力，扩展性。因此，C4 是高度模块化的系统，除了必需的子系统外，很方便扩充。图 4-1 是 C4 训练的一只“牧羊狗”。



图 4-1

C4 与世界之间的相互作用通过数据记录(DataRecord)的形式传输，数据记录可以表示 C4 的感应器和感知系统所能识别的任何信息，包括声音、视觉、符号等。

图 4-2 是 C4 脑的体系框图，每一个部分都可以通过端口对黑板进行读写访问。

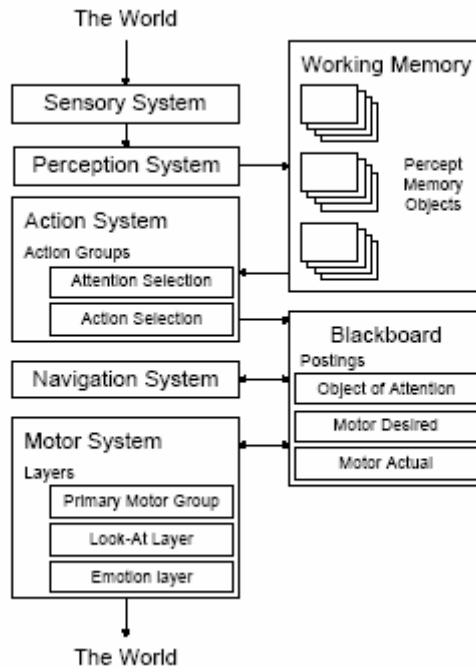


图 4-2

4.1.1 感应系统 Sensor System

为了保持世界——Agent 分离的完整性,C4 采用抽象感应系统(Sensor System Abstraction)。感应系统有两个作用：

- (1) 过滤器，比如牧羊狗不应该看到她后面的羊；
- (2) 转换器，因为同一条信息在不同的 Agent 眼里意义并不相同。

4.1.2 感知系统 Perception System

感知系统将“意义(Meaning)”赋予感应到的事件。感知子(Percept)是感知系统的原子类型，又是一种数据提取单元。感知子的输入是从感应系统传过来的数据记录，返回匹配概率(SheepShapePercept 返回看到的是一头羊的概率)，如果概率大于阈值，同时返回提取的数据(SheepShapePercept 羊的位置坐标)。C4 没有对感知子的实现方式做任何假设，因此开发者可以根据需求选择算法，这是模块化增强可扩展性的直接体现。

为了提高感知效率，感知子以树形结构(图 4-3)组织，每一次判断都裁减掉大量不可能的感知判断，降低复杂度；而且子感知子只接受父感知子提取的数据；减少数据量。如果一个声音信号直接传递给 Sound 感知子，而不再考虑 Shape 等感知子。每当数据记录到达，感知系统就会创建一个感知存储(PerceptMemory)对象，此后每一个感知子返回的可信度和提取的数据都保存在感知存储中。感知

存储代表 Agent 对刺激理解的所有内容。

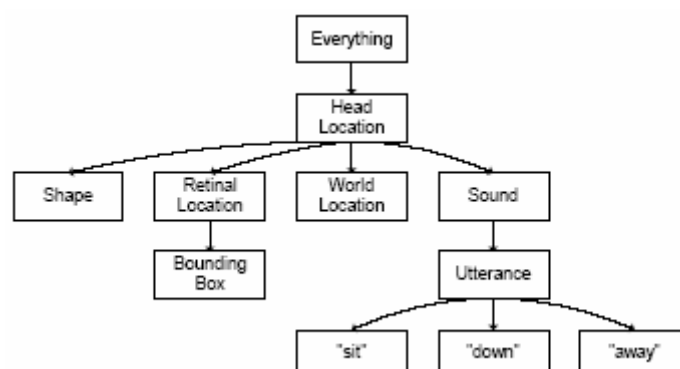


图 4-3

感知子不仅仅作为接受场(Receptive Field)，而且会根据反应系统(Action System)的反馈，修改树的拓扑结构，动态添加子节点——称之为革新(Innovation)。

4.1.3 工作存储 Working Memory

C4 包含一个 Working Memory 结构，储藏驻存的 Percept Memory 对象，保存对世界中的物体的感知历史。她们构成生物对世界的“视图(View)”。

4.1.4 预测和惊奇 Predication and Surprise

基于 Percept Memory 对象，生物可以模仿下意识，对现在或将来的状态作出预测。预测一般对感知存储的数据进行外推函数逼近，更复杂地考虑周期性为和知行关系。预测偶尔出现的偏差，可以作为惊奇的基础。

4.1.5 动作系统 Action System

动作是一个五元组 ActionTuple:

- (1) Primitive Action(s): 做什么;
- (2) TriggerContext: 什么时候做;
- (3) ObjectContext: 对谁做;
- (4) DoUntilContext: 做多久;
- (5) Intrinsic Value: 做的好处。

C4 有 Attention Group 和 Primary ActionGroup 两个动作组(ActionGroups)，每个动作组有两条 ActionTuple 链表: Startle、Default。Startle 具有较高的优先级，具有最高优先级的 ActionTuple 被激活; Default 根据 e-values 随机的激活动作。

学习是 C4 的研究重点，类似人类训练狗的方法，有三种学习方式: Credit

Assignment, State-space Discovery, Innovation。

4.1.6 导航系统 Navigation System

导航系统是对动作系统的高级实现，她隐藏动作的实现细节，使动作简化成“前进、吃、离开”等一般的指令。

导航系统从动作系统接受指令，将之分解优化成恰当的运动命令，再发送给运动系统。

4.1.7 运动系统 Motor System

运动系统组织成动词图(Verb Graph)，定义了最基本的运动和运动序列。运动节点是手工设置的，每个节点可以由副词空间(Adverb Space)描述，如向左、笔直、向右地行走。参数从黑板的 MOTOR_DESIRED 和 MOTOR_ADVERB 获得。

分层使生物能同步进行多个运动。如牧羊狗的头部层、身体层、尾巴层。

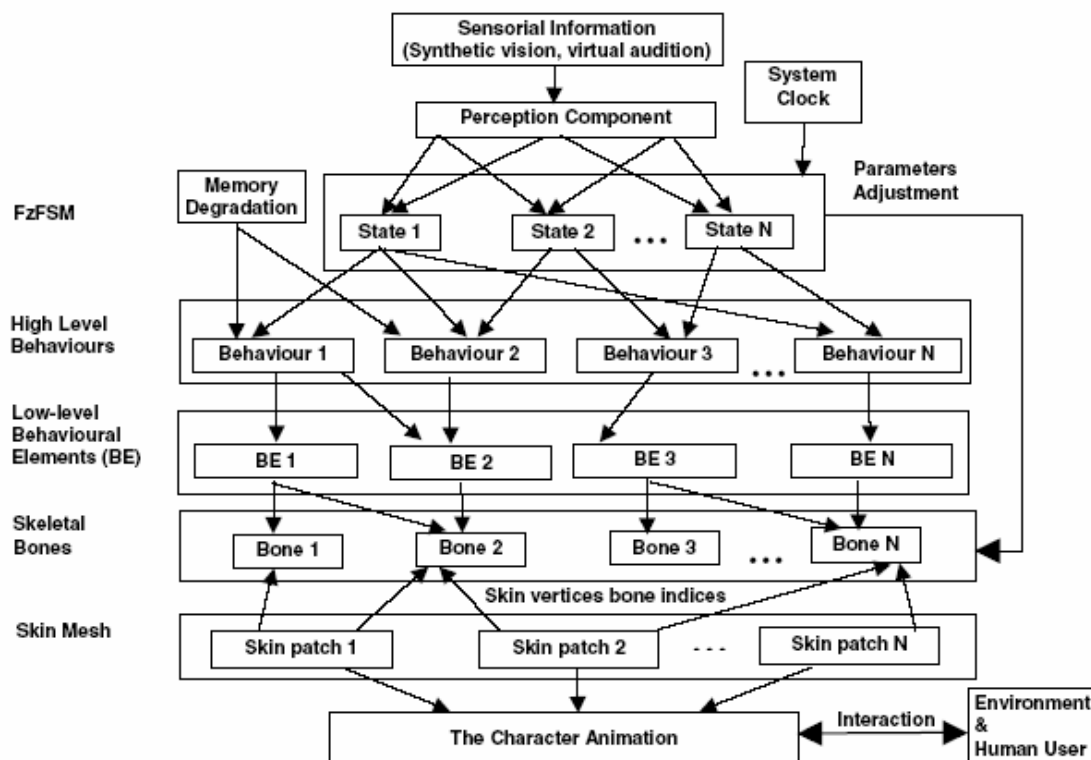


图 4-4

4.2 一种可视化的新方法 A New Approach for Visualizing

Wen 的方法主要有三个部分：视觉合成(Synthetic Vision)，记忆 FzFSM，动画库；图 4-4。

4.2.1 感觉输入 Sensorial Input

系统有两个信息通道：视觉和听觉，主要是视觉。有别于一般的直接感应，

一种视觉合成方法——多层次空间场景，用于表示外部世界。3D 的虚拟环境由多层次空间表示成八叉树(Octree)，如果一个节点包围的图元大于某个阈值，就将这个节点再次叉分成八个子结点。视觉合成用八叉树结合局部 Z 缓冲算法，判断 Agent 的可见物体。同样地，感知层也要对输入进行过滤。

4.2.2 FzFSM 和记忆 FzFSM and Memory

FzFSM 是一个有限状态机，每一个状态有介于[0,1]的模糊值代表状态等级，调节动画的转移矩阵，避免重复的动画。

角色的行为设计分成高级和低级两层。高级层面向目标，更概括，更直观。低级层面向动作，更具体，更细致。如一个“追逐猎物”的高级行为可分解成“直冲”，“转弯”，“加速”等低级行为。

记忆就是将能记住的物体保留在八叉树中，并用一个递减的八进制数说明记忆的准确度。

4.2.3 动画层 Animation Layer

动画层用面片皮肤动画(Mesh Skinning Animation)实现，激活骨骼(Skeleton)尾部的转移矩阵完成动画。Quaternion 方法的球面线性插值(Spherical Linear Interpolation)算法完成动画的平滑过渡。

图 4-3 分别表示行走的人，边走边小心观望的人，惊恐的跑的人。



图 4-3

4.3 小结

行为 Agent 主要从行为上模拟生物，代替人类在虚拟环境中活动，给其它研究一个相对理想的试验平台，不论是动力学、社会学、生物学。

行为 Agent 与特定的硬件结合，还可以帮助人类完成繁重危险的工作。如汽车装配机器，外太空探测器等。

然而，现在的研究还不足以令行为 Agent 让人信服，只是在特定的条件下才能发挥特点。如让 Agent 平衡自然上下楼梯就不是一件易为的工作。

我想行为 Agent 发展的将来就是机器人,与人类一起生活工作在现实世界中,距完全实现这一天也许很远,但在一定程度上获得突破很值得期待。

五 多 Agent 系统 Multi-Agent

在大型的虚拟环境中,一个 Agent 并不能很真实的反映客观世界的情况,常常需要大群 Agent 参予其中,相互交流,协同合作。比如社会系统的模拟, [35]

称之为 MAS(Multi-Agent Society); [36]中蚂蚁王国的例子是基于计算 Agent (Computational Agent), CA 代表一一对应关系的有机体。

以下以金斯敦大学计算和信息学院的 FreeWill[33][34]体系构架为例,说明多 Agent 在生成人群场景上的应用。

FreeWill 的目的是利用多 Agent 完成人群场景的模拟,设计实现一套可扩展可调节的认知构架,并和其它软件包协同工作,如几何引擎, AI 引擎等。FreeWill 用 java 编码,图形包用的是 3D Max Studio。

5.1 虚拟人 Avatar

FreeWill 的虚拟人集合 FCA[Funge’s Cognitive Architecture]和 SAC[32]的特点,并扩展成更鲁棒更灵活的体系。表 5-1 是关于 FCA 和 SAC 的对比。

Architecture name	FCA	SAC
Architecture orientation	Character oriented	Agent oriented
Recognition for the ani- mation aspect	Yes	No
Representation of the character’s knowledge	Internal world model + domain knowledge	Beliefs (a knowledge base)
Autonomy	Yes	Yes
Learning	Recognition for learning	No explicit emphasis on learning
Reactive and pro-active behaviour	Yes	Yes
Interaction with other characters	Implicit	Explicit
Main implementation cy- cle	Sense-think-act	Sense-think-act
Sensing/perceiving	Updates the world model	Gives rise to (internal) events
Planning	Search on atomic actions	Plan libraries
Ease of implementation	Not obvious	Easy to follow develop- ment model
Goals	Yes	Yes
Formalism	Situation calculus	Under development

表 5-1

每个虚拟人都是改进的 SAC Agent, 有两个子系统: 几何引擎和 AI 引擎。

几何引擎包括一系列几何函数, 把握虚拟人的躯体参数。几何引擎是从 FCA 体系中借鉴过来, 以反运动学控制虚拟人的几何面。几何引擎内建了许多简单的动作模式模拟人的行为。

AI 引擎完成环境感应, 数据感知, 目标规划, 碰撞避免等。FCA 的内部世

界模型是知识的基本组成，同样包含计划库(Plan Library)。

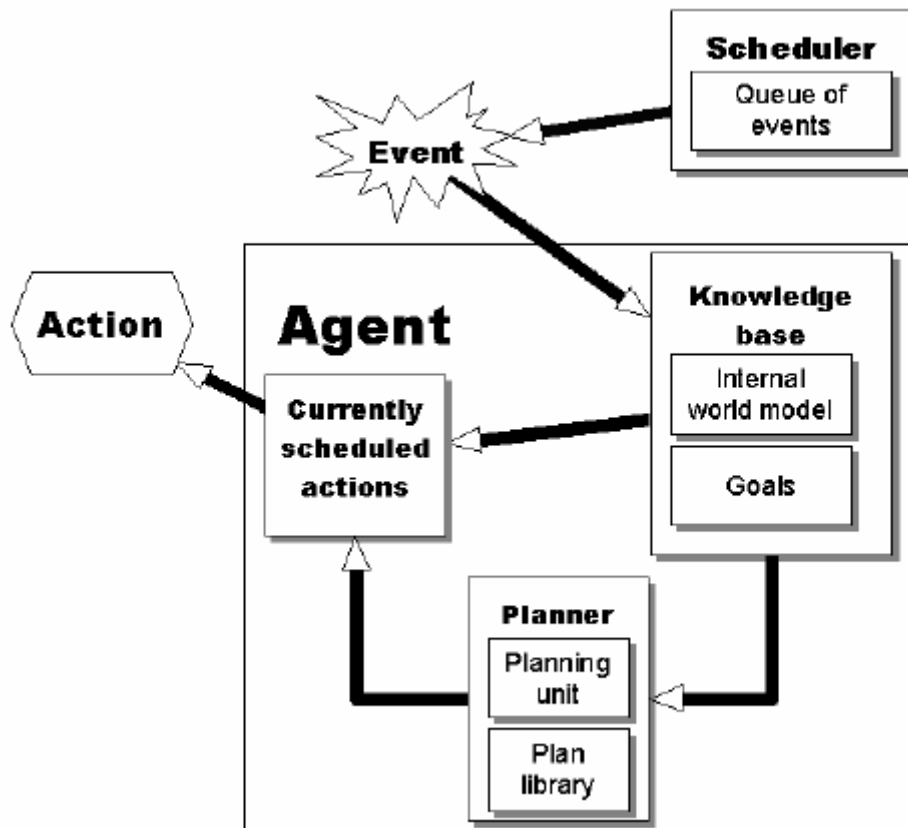


图 5-2

5.2 虚拟人间的通信 Communication among Agents

虚拟人之间的通信主要通过基于事件(Event)的全局调度(图 5-2)实现。调度器(Scheduler)维护一个事件队列，既分配事件，又要安排事件的序列。虚拟人必须先提交一个事件到调度器，等到接受这个事件才能执行相应的命令。

所以 FreeWill 执行周期是这样进行的：

- (1) 虚拟人感应环境的变化；
- (2) 更新信念(Belief)；
- (3) 评估当前计划，如果一个新的动作必须被执行，取消当前运动队列，提交新的动作事件；
- (4) 如果需要，更新目标；
- (5) 如果需要，更新计划；
- (6) 选择上一个动作并提交，当虚拟人执行一个动作时，也会提交一个感应事件。

例如两个虚拟人都要走过环境的某一个位置，感应时都发现没有其他人，于是提交行走动作，此时，调度器就可以延迟其中一个动作事件的分配，使一个虚拟人等待。

目的层次切换也是虚拟人交互的重要手段。起初，要为每一个虚拟人设定一个主目的，然后驱动她们。在运动过程中，虚拟人根据内部知识和感应数据，判断是否有新目的，并会切换执行。比如，虚拟人在走到一点的过程中遇到一个朋友，会将主目的替换成第二目的，重新计划同步握手动作，之后如果没有其它意外，又切回主目的，继续行走。

图 5-3 是 FreeWill 模拟人行道上行人们的结果。



图 5-3

5.3 小结

相对单个 Agent，多 Agent 替社会系统等群体结构的研究打下了基础。然而体系构架更为复杂，现有的进展尚且不能满足人们的实践要求。

将来的发展应该有：

- (1) 交流基础上的协作，如 FreeWill 系统的虚拟人就不能合作抬起一个箱子；

- (2) 个体和群体相结合，不仅仅在整体结构上模拟，而且单个 Agent 也具有多种多样的生物行为；
- (3) 环境更加真实，Agent 也更加适应，如果给环境加入天气、灾难、谣言等因素，Agent 也应该可以模拟人的行为。

六 总结 Conclusion

通过以上的阅读学习，对智能 Agent 在虚拟环境中的应用，我形成了一个大致框架。智能 Agent 这个概念虽然在很多应用程序中都存在，但是真正能提出一个完整的体系框架的系统却寥寥无几。因此，智能 Agent 其实并没有在市场上

普及。

随着人们对计算机越来越熟悉，自然不会再满足于单方向的控制，不论是商业还是个人的应用，都希望计算机更聪明，能自主地选择具体命令，合理协调规划，不仅正确而且高效地完成规定的目标。同时，给计算机一个人性化的外表，而不仅是一堆冰冷的金属，也是家庭应用的发展方向。这将促进多通道、多方位的人机交流。

我想，智能 Agent 应该是一个与现实世界的性质较为吻合的理论，偏向于实际，可以很容易地引入到系统中，当如果设计不当，勉强使用，并不能起到应有的效果，反而混淆概念，仅给系统披上人工智能的流行外衣，内里是一塌糊涂。

从所查阅的论文看出，美国的研究者自然投入很多资源，而欧洲和日本的大学和研究机构在智能 Agent 科研上也作了很多工作，取得显著的成绩。

七 引文 References

以下引文全部来自互联网，仅仅仔细阅读其中的一部分，其它只是粗略的浏览归类。

[1] Stan Franklin , Art Graesser, Is it an Agent, or just a Program?:A Taxonomy for

Autonomous Agents, Institute for Intelligent Systems, University of Memphis, Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag, 1996.

[2] *Helmut Prendinger, Santi Saeyor, and Mitsuru Ishizuka*, Animated Agents for Language Conversation Training, Department of Information and Communication Engineering, School of Engineering, University of Tokyo

[3] *Ron Cole, Tim Carmell, Pam Connors, Mike Macon, Johan Wouters, Jacques de Villiers, Alice Tarachow*, Intelligent Animated Agents for Interactive Language Training, Center for Spoken Language Understanding, Oregon Graduate Institute, Portland, Oregon

[4] *Marc Evers, Anton Nijholt, Jacob*: An Educational Agent in a Virtual Environment, Centre of Telematics and Information Technology (CTIT) University of Twente

[5] *Shoham*, “Agent-oriented Programming”

[6] *Ipke Wachsmuth, Yong Cao*, Interactive Graphics Design with Situated Agents, Faculty of Technology, University of Bielefeld

[7] *Norman Badler*, Virtual Humans for Animation, Ergonomics, and Simulation, Center for Human Modeling and Simulation University of Pennsylvania Philadelphia

[8] *Elisabeth André, Thomas Rist*, Controlling the Behaviour of Animated Presentation Agents in the Interface:Scripting versus Instructing

[9] *Helmut Prendinger, Mitsuru Ishizuka*, SCREAM: Scripting Emotion- based Agent Minds, Dept. of Information and Communication Eng. Graduate School of Information Science and Technology, University of Tokyo

[10] *Yasmine Arafa, Kaveh Kamyab, Ebrahim Mamdani, Sumedha Kshirsagar, Nadia Magnenat-Thalmann, Anthony Guye-Vuillème, Daniel Thalmann*, Two approaches to Scripting Character Animation

[11] *Johan F. Hoorn , Anton Eliëns , Zhisheng Huang , Henriette C. van Vugt, Elly A. Konijn, Cees T. Visser*, Agents with Character: Evaluation of Empathic Agents in Digital Dossiers

[12] Animated Pedagogical Agents in Multimedia Educational Environments: Effects of Agent Properties, Picture Features, and Redundancy

[13] **Cholyeun Hongpaisanwiwat*, ***Michael Lewis*, Attentional Effect of Animated Character, *Department of Computer Science, Thammasat University, **School of

Information Sciences, University of Pittsburgh

[14] *Stan Franklin, Art Graesser, Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents* Institute for Intelligent Systems University of Memphis Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag, 1996

[15] **Michael Wooldridge*, ***Nicholas R. Jennings*, *Intelligent Agents: Theory and Practice*, *Department of Computing Manchester Metropolitan University, **Department of Electronic Engineering Queen Mary & Westfield College, Knowledge Engineering Review, October 1994

[16] *Stuart Russell, Peter Norvig, Artificial Intelligence: A Modern Approach*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2003

[17] *Ian Paul Wright*, *Emotional Agent*, School of Computer Science, Cognitive Science Research Centre, University of Birmingham, A thesis submitted to the Faculty of Science of the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

[18] **Yukiko I. Nakano*, ***Masashi Okamoto*, ***Daisuke Kawahara*, ***Qing Li*, ***Toyooki Nishida*, *Converting Text into Agent Animations: Assigning Gestures to Text*, *Japan Science and Technology Agency, **The University of Tokyo

[19] *Dominic W. Massaro, Michael M. Cohen, Jonas Beskow, Sharon Daniel, and Ronald A. Cole*, *Developing and Evaluating Conversational Agents* Perceptual Science Laboratory University of California

[20] *Stefan Kopp, Ipke Wachsmuth*, *Model-based Animation of Coverbal Gesture*, Artificial Intelligence Group Faculty of Technology, University of Bielefeld

[21] *Christian Becker, Stefan Kopp, and Ipke Wachsmuth*, *Simulating the emotion dynamics of a multimodal conversational agent*, Artificial Intelligence Group Faculty of Technology, University of Bielefeld

[22] *Multimodal Dialogs with a Conversational Agent*

[23] *Jonas Beskow*, *ANIMATION OF TALKING AGENTS*, Centre for Speech Technology KTH, 1997

[24] *Norman Badler, Jan Allbeck, Liwei Zhao, Meeran Byun*, *Representing and Parameterizing Agent Behaviors*, Center for Human Modeling and Simulation, Computer and Information Science, University of Pennsylvania

[25] *Vinoba Vinayagamoorthy, Mel Slater, Anthony Steed*, Emotional personification of humanoids in Immersive Virtual Environments, Department of Computer Science, University College London

[26] *徐晓媛(Xiaoyuan)*, Artificial Animals for Computer Animation: Biomechanics, Locomotion, Perception, and Behavior, Department of Computer Science, University of Toronto, for the degree of Doctor of Philosophy, 1996

[27] *James J. Kuerner, Jr*, AUTONOMOUS AGENTS FOR REAL-TIME ANIMATION, the department of computer science of stanford university, for the degree of Doctor of Philosophy, December 1999

[28] *Armin Bruderlin, Sidney Fels, Silvio Esser, Kenji Mase*, Hierarchical Agent Interface for Animation, ATR Media Integration & Communication, Japan

[29] *Toni Conde, William Tambellini, and Daniel Thalmann*, Behavioral Animation of Autonomous Virtual Agents Helped by Reinforcement Learning, Virtual Reality Lab, Swiss Federal Institute of Technology (EPFL)

[30] *John P. Granieri, Welton Becket, Barry D. Reich, Jonathan Crabtree, Norman I. Badler*, Behavioral Control for Real-Time Simulated Human Agents, Center for Human Modeling and Simulation, University of Pennsylvania

[31] *Damian Isla, Robert Burke, Marc Downie, Bruce Blumberg*, A Layered Brain Architecture for Synthetic Creatures, The Media Laboratory Massachusetts Institute of Technology, pp. 1051-1058, in Proceedings of Seventeenth Joint Conference on Artificial Conference IJCAI-01, 4-10 August, Seattle, USA, 2001

[32] *Zhigang Wen, Q.H. Mehdi, N.E. Gough*, A New Animation Approach for Visualizing Intelligent Agent Behaviours in a Virtual Environment, School of Computing and Information Technology University of Wolverhampton

[33] *Michael Winikoff, Lin Padgham, and James Harland*, Simplifying the Development of Intelligent Agents, in AI2001: Advances in Artificial Intelligence, 14th Australian Joint Conference on Artificial Intelligence, LNAI 2256, pp 557-568, Adelaide, December 2001

[34] *Adam Szarowicz¹, Peter Forte²*, Combining Intelligent Agents and Animation, ¹WhiteSpace Studio, Kingston University, ²School of Computing and Information Systems, Kingston University

[35] *Adam Szarowicz¹, Peter Forte², Juan Amiguet-Vercher¹, Petros Gelepithis²*, Application of Autonomous Agents for Crowd Scene

Generation, ¹WhiteSpace Studio, Kingston University, ²School of Computing and Information Systems, Kingston University

[36] *Alexander Artikis, Kloyd Kamara, Frank Guerin, Jeremy Pitt*, Animation of Open Societies, Imperial College of Science, Technology and Medicine, Electrical & Electronic Engineering, Intelligent & Interactive Systems Group

[37] *Alexis Drogoul, Jacques Ferber*, Multi-Agent Simulation as a Tool for Modeling Societies: Application to Social Differentiation in Ant Colonies, LAFORIA, Boîte 169, Université Paris VI, 75252 PARIS CEDEX 05 FRANCE