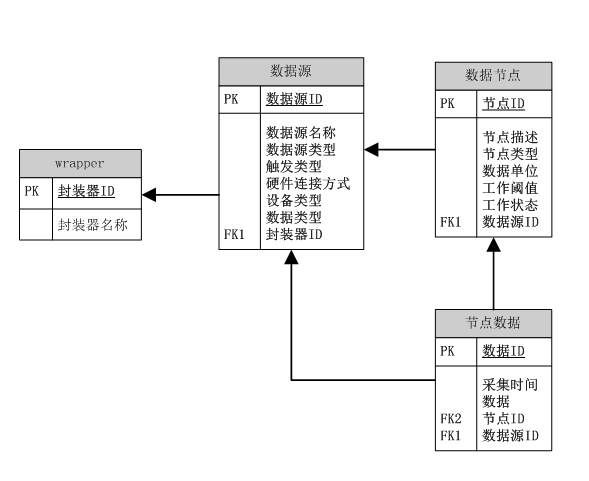
# 第4章 基于CA-DHS的残疾人健康服务上下文感知技术

前面着重探讨了残疾人健康服务的各项需求，并在几种经典的上下文感知系统框架基础上设计了适合于智慧社区中残疾人的系统框架。本章将分别就系统框架中感知层、上下文处理层和应用层的重要技术细节展开论述。

## 4.1 原始上下文信息的采集与封装

由于采集的底层上下文信息格式各不相同，如血压、心率等生理信息、以坐标形式反映的GPS定位信息和以实际地点表示的RFID定位信息等，因此这里设计的传感设备wrapper提供一个统一的接口，以实现对传感器数据和外部系统的封装，使系统上层可以屏蔽底层硬件设备与数据格式的差异。作为服务的最小单元，每一个用户（残疾人）都对应着一个wrapper，其中封装了位置与生理状态两类数据。我们将这类数据定义为数据源，每个数据源又包括了若干数据节点，即基本的传感设备。数据源和数据节点都具有统一的描述格式，组成了wrapper的基本结构。图4.1是数据源、数据节点以及节点所采集数据的数据库模型图，用来描述wrapper的主要组成结构。

**图4.1 传感设备wrapper的数据库模型图**

在残疾人健康服务系统中，这些数据的采集是周期性不断循环进行的，以实现对残疾人健康状况的持续监控。传感设备wrapper向上层提供了数据访问接口，以统一的节点数据的形式为上下文的进一步处理屏蔽了硬件细节。此外，wrapper还封装了传感设备的控制接口，使系统能够根据需要控制传感器的上传频率、采样频率。考虑到底层的数据采集速度与上层应用使用数据的速度可能存在不一致、不匹配的问题，这里还设计了一个上下文数据缓存。数据缓存采用循环队列的数据结构，申请存储空间的操作只需要进行一次，在采集与使用上下文数据时，无需再次申请和释放资源。传感设备wrapper总是可以把最近得到的数据推送到上下文数据缓存中。

通过底层传感设备采集的上下文信息分为位置信息和生理状态信息两类，即由用户随身携带的各种传感器获取的残疾人血压（Bp）、心率（Hr）、呼吸频率（Br）及血氧饱和度（SpO2）等四种生理状态信息和RFID阅读器读取标签与GPS定位结合使用获取到的位置信息（Position）。这些数据信息都不能被推理引擎所直接识别，因此论文使用了七元组DHS\_Context =（ID，position，bp，hr，br，SpO2，time）的方法来对采集的原始上下文进行形式化描述，其中ID代表用户信息，time代表数据采集时间。

例如，残疾人小张(ID是CA0006)下午15:00正在社区的20栋401，其血压为113/80mmHg、心率85次/分、呼吸频率17次/分、血氧饱和度为98%，那么根据以上对各项生理状态的离散化处理，我们就可以用一个七元组来描述这一信息：（CA0006，20-401，bp-regular，hr-regular，br-regular，SpO2-regular，15:00）。

论文中是通过读取DHS\_Context七元组中经过解释的上下文信息，来实现残疾人健康服务上下文的表示与建模的。

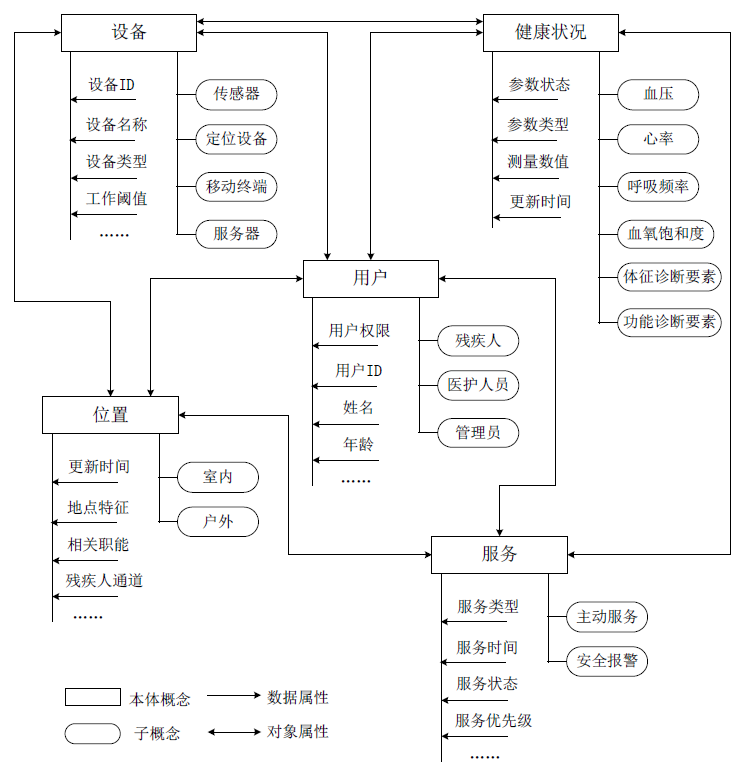
## 4.2 残疾人健康服务上下文的表示与建模

由于结构和流程的复杂性，上下文感知计算系统在设计开发以及测试部署等方面都具有较高的难度。因此，采用建模、仿真的方法不仅能独立地研究上下文感知计算所涉及的各个领域，解决实际应用中遇到的诸多困难，同时还能降低项目风险，节约成本。

本体模型是目前被上下文感知技术广泛采用的一种建模方法。本体的方法有多种，而OWL(Web Ontology Language)作为W3C极力推荐的一种标准的本体描述语言，建立在RDF、RDFS以及DAML+OIL的基础上，能够清楚地表达程序中需要处理的文件中的知识。

基于上下文感知的残疾人健康服务要求实现对残疾人健康状况及位置的监控，提供主动的个性化服务及预警，因此我们在系统中设计了五个本体，分别是用户、设备、健康状况、位置以及服务。其中，以用户本体为中心，其他四个本体都与用户存在一定的关系，这也体现了普适计算以人为中心的思想，其最终目的就是为残疾人提供更加智能、更加方便的健康服务。同时，建立这些本体还需要读取DHS-Context七元组中的上下文信息，以将用户和其位置、健康状况对应起来。

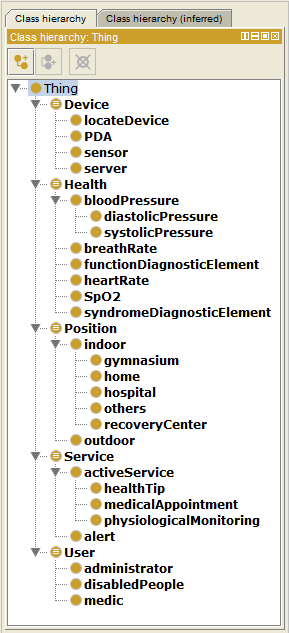
图4.3描述了系统中各本体之间的关系及各自的属性。



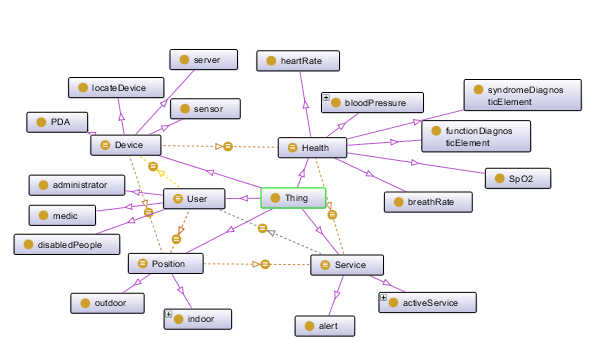
**图4.3 残疾人健康服务系统本体结构图**

有了本体之间的关系及各自的属性，就可以通过Protégé[本体](http://zh.wikipedia.org/wiki/%E6%9C%AC%E4%BD%93%E8%AE%BA_(%E8%AE%A1%E7%AE%97%E6%9C%BA))编辑器来本体类和相应的本体结构图。

使用protégé创建的本体类如图4.4所示。



**图4.4 使用protégé创建的本体类**

图4.5是protégé根据定义的本体类生成的本体结构图。

**图4.5 protégé生成的本体结构图**

在图4.4和图4.5中可以看到，Thing类是所有其他类的公共父类，其下的User、Health、Service、Position、Device类除了各自具有自己的子类以外，其相互之间还存在着各种属性，这些属性体现出使用本体方法来表示上下文的优势。以下是部分描述残疾人本体模型的OWL代码:

<owl:Class rdf:ID=“Indoor”>

<rdfs:subClassOf>

<owl:Class rdf:ID=“Position”/>

</rdfs:subClassOf>

</owl:Class>

<owl:Class rdf:ID=“Home”>

<rdfs:subClassOf>

<owl:Class rdf:ID=“Indoor”/>

</rdfs:subClassOf>

</owl:Class>

<owl:ObjectProperty rdf:ID=“located\_in”>

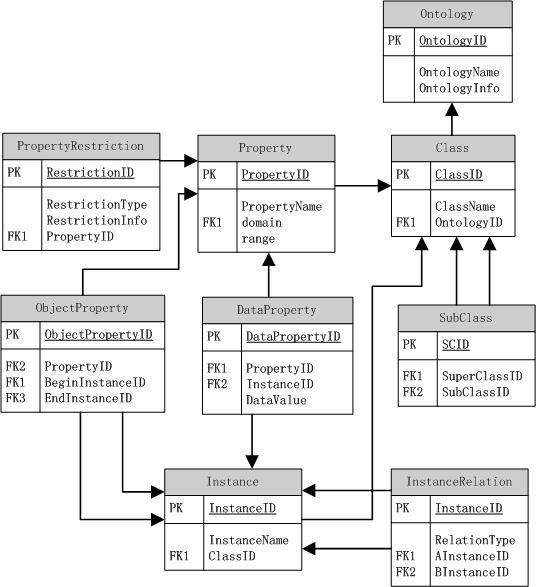
<rdfs:domain rdf:resource=“#User”>

<rdfs:range rdf:resource=“#Position”>

</owl:ObjectProperty>

## 4.3 残疾人健康服务上下文的存储与管理

由于CA-DHS框架中以本体来完成上下文的建模，因此我们在存储残疾人健康服务上下文时需要对本体的存储方法进行研究。在本系统中，我们采用了关系数据库来对本体进行存储。在关系数据库中存储本体应当具有较高的规范化程度、易于理解和查询、稳定的结构以及完备的信息等。遵照这些原则，我们设计了残疾人健康服务上下文本体的在关系数据库下的存储模式。如图4.6所示。



**图4.6 本体在关系数据库下的存储模式**

Class表用于描述类，类作为本体模型中的基本组成部分，与属性、实例及约束都有着密切的联系，该表中定义了类的ID、名称以及所属本体的ID。SubClass表、Ontology表只与Class表发生直接关联，其中SubClass表保存Class之间的从属关系，Ontology表则记录本体的信息。

在属性表Property中，domain描述该属性所属的类，而range则描述该属性的取值范围。由于OWL将本体的属性划分为ObjectProperty和DataProperty，它们各自有着不同的描述方式，因此我们将这两种属性分别保存到 ObjectProperty表和DataProperty表中。同时，我们还为一些使用频率很低的约束属性设计了一张PropertyRistriction表，该表保存了类someValuesFrom、allValuesFrom等约束，但是只使用了一张表进行保存，目的是降低数据库表的数量，以便对数据库进行管理。

对实例的保存使用Instance表，以其对应的ClassID作为外键，同时用InstanceRelation表存储实例之间的关系，一般是指相同实例或者不同实例的关系。

按照这种方式在关系数据库中存储本体，能够满足各种常用的查询要求，如查询某个类的直接父类或全部直接子类、某个类的所有属性、某个类的全部实例，或者按条件查找本体中时候包含某个概念。对于查询频率较低的信息，由于设计了不同的数据库表，这样就提高了查询的效率。同时，这种存储模式结构简单、稳定便于修改和维护。

## 4.4 残疾人健康服务上下文的推理

在CA-DHS中，我们将残疾人健康服务上下文的推理分为生理状态推理、本体推理和规则推理。其中生理状态推理用于根据采集到的几种生理状态数据来判断残疾人是否健康，以此作为进一步推理或报警等操作的依据；本体推理是从构建的本体中根据对象属性(ObjectProperty)的逻辑关系进行的简单推理，比如某人在病房，而病房是医院的一部分，则推断某人在医院中；规则推理则是结合各类上下文并根据规则库中定义的推理规则，完成各种不同的服务。

在系统开发过程中，根据实际需要，我们采用protégé本体建模工具来创建和编辑本体，通过Netica Java API来构造生理状态信息的贝叶斯网络并进行推理，利用Jena Java开发工具包来对建立的残疾人本体模型进行基于规则的推理。

### 4.4.1 基于贝叶斯网络的生理状态上下文推理

论文在对残疾人的血压、心率等几项生理状态上下文进行处理时，引入了贝叶斯网络的方法，它以贝叶斯概率理论为基础，被广泛用于不确定性上下文的推理中。

朴素贝叶斯（naïve Bayes）理论是基于贝叶斯定理与特征条件独立假设的分类方法。对于给定的训练数据集，首先基于特征条件独立假设学习输入/输出的联合概率分布；然后基于此模型，对给定的输入x，利用贝叶斯定理求出后验概率最大的输出y。

设输入空间χRn 为n维向量的集合，输出空间为类标记集合Υ={c1，c2,…,ck}.输入为特征向量x∈χ，输出为类标记（class label）y∈Υ。X是定义在输入空间χ上的随机向量，Y是定义在输出空间Υ上的随机变量。

先验概率分布如公式（4-2）。

（4-2）

条件概率分布由公式（4-3）定义，即联合概率分布。

（4-3）

后验概率计算根据贝叶斯定理进行，如公式（4-4）。

 （4-4）

结合以上条件概率分布，可以得出公式（4-5）。

 （4-5）

公式（4-5）反映了朴素贝叶斯理论所采用的原理，即后验概率的最大化。

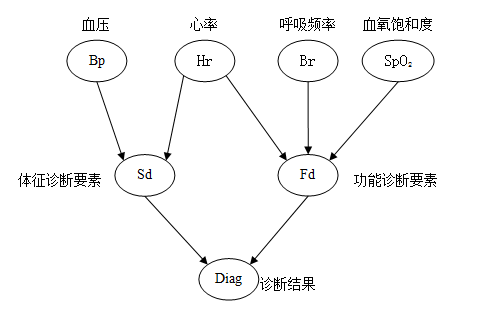
朴素贝叶斯理论能够假设输入变量都是条件独立的，如果假设它们之间存在概率依存关系，模型就变成了贝叶斯网络。

一个贝叶斯网络由两部分组成，有向无环图（Directed Acyclic Graph ,DAG）和条件概率表（Conditional Probability Table，CPT），其形式化定义如公式（4-6）。

（4-6）

在以上定义中，<N,E>代表有向DAG，N是图中节点的集合，E则表示连接节点的有向边的集合；θ代表CPT。

结合领域专家的指导和医疗知识，我们构造了残疾人生理状态上下文推理的贝叶斯网络的有向无环图，如图4.7所示。

**图4.7 生理状态的贝叶斯网络图**

此外，根据网络中各节点的概率分布，我们以专家知识和统计数据拟定了条件概率表（CPT）。该节点的条件概率共有27种组合，由于篇幅所限，仅在表4.5给出其中一部分。

**表4.5 功能诊断要素的CPT（条件概率表）**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hr** | **Br** | **SpO**2 | **P(Fd=abnormal)** | **P(Fd=dangerous)** |
| regular | regular | regular | **0.95** | **0.05** |
| abnormal | regular | regular | **0.75** | **0.25** |
| dangerous | regular | regular | **0.4** | **0.6** |
| regular | abnormal | regular | **0.8** | **0.2** |
| abnormal | abnormal | regular | **0.6** | **0.4** |
| **……** | | | | |
| regular | dangerous | dangerous | **0.25** | **0.75** |
| abnormal | dangerous | dangerous | **0.1** | **0.9** |
| dangerous | dangerous | dangerous | **0.01** | **0.99** |

四项基本生理指标都按照上文所述分类为regular、abnormal和dangerous，诊断要素和诊断结果则分成regular和dangerous两类。四项基本生理指标的先验概率都应当是该指标出现异常的一般概率，在大量统计的基础上也可以认为是该项生理指标出现异常的人群占社会总人口的比例。

论文将采集的几种基本生理状态参数均按照其一般发病率来定义先验概率，以Bp（血压）节点为例，根据近年来统计数据，我国高血压的发病率为18.8%，恶性高血压占到其中的27.3%，因此Bp的先验概率可以预设为：P（Bp=regular）=0.81，P（Bp=abnormal）=0.14，P（Bp=dangerous）=0.05。

论文构造的贝叶斯网络，经过推理后返回的是最终诊断结果diag的概率值，该值的大小反映了残疾人的健康状况。经过大量的模拟实验，对最终诊断结果Diag的概率取值范围做出如下定义：

若，则残疾人状态良好，不做任何处理；

若，则残疾人状态较差，做出相应的提示；

若，则残疾人处于危险状态，急需救助。

**4.4.1.4 贝叶斯网络的推理**

Jena提供了一个解析函数，能够把本体中的概念及实例、先验概率、条件概率等信息解析出来，作为贝叶斯网络的输入条件，再调用Netica java API来构造贝叶斯网络并实现推理。基于Netica的贝叶斯网络构建和推理过程如图4.8所示。

**图4.8 基于Netica的贝叶斯网络构建和推理过程**

先验概率

概念及实例

条件概率

格式转换

本体映射为贝叶斯网络

贝叶斯网络推理

概念及实例

条件概率

概念关系

CPT

CPT

构建节点

构建有向边

Netica

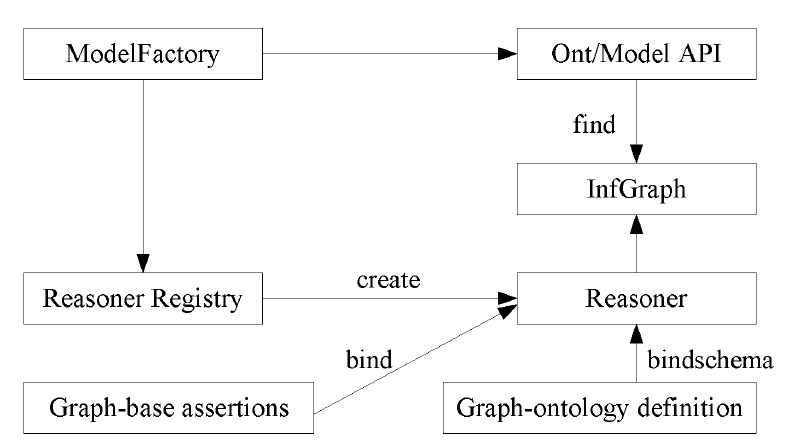
Netiva Java API支持联合树推理算法。因策先将贝叶斯网络的有向图结构转换为联合树，再将贝叶斯网络中的条件概率表分配到联合树的各簇节点，并对它们进行初始化，得到满足全局一致性的联合树。再在构建的联合树中进行消息传递，求出各随机变量的条件概率，完成贝叶斯网络的推理。论文设计了一个仿真实验，实验条件为：采集到的生理信息数据分别为血压117/85mmHg、心率105次/分、呼吸频率22次/分、血氧饱和度为97%。实验在MyEclipse环境下利用Netica完成。

经过数据的离散化处理，得到Bp= regular，Hr=abnormal，Br= abnormal，SpO2 = regular。将这些数据作为推理证据，加载到贝叶斯网络中进行推理，使用Netiva Java API提供的联合树推理算法，经过计算得到Diag的后验概率为0.4335。

仿真结果P(Diag)∈（0.25,0.45]，此时应该为残疾人进行提示，使残疾人能根据自身状况做出自我调整。即在残疾人心率和呼吸频率高于正常水平时，判断残疾人处于异常状态。事实上，该状态下的人心慌、头晕、呼吸急促等症状，可能是由于某些心脏疾病轻度发作所导致的，根据主动提示，残疾人可以选择服用药物或者调整呼吸等手段进行自我处理。

### 4.4.2 基于Jena的本体推理及规则推理

Jena是由惠普公司开发为语义网（Semantic Web）研究项目开发的一个开源Java语义工具包，为OWL、RDF、RDFS等本体描述语言提供了一个程序开发环境，并带有对本体进行解析、推理、查询等操作的函数调用与接口[42]。

****Jena的推理机是基于规则构造的结构如图4.12所示。

**图4.12 Jena的推理机结构**

Jena实现本体推理，是借助本体模型内置的对象属性（ObjectProperty）来进行的。

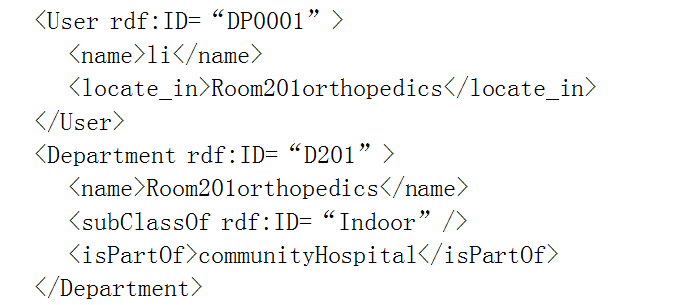
(1)本体推理

下面通过一个实例来描述上下文的上下文知识库的本体推理过程。

系统在读取采集到的上下文信息后，会以本体的形式保存在自动生成的DisabledService.owl文件中。现在系统获取到ID为DP0001的用户小李正位于201房间骨科，我们在DisabledService.owl文件中能够找到生成的相应代码，截取的一段owl代码如下所示：

Jena读取DisabledService.owl，获得上述信息后进行本体推理。论文通过Jena进行本体推理的部分实现代码如下：





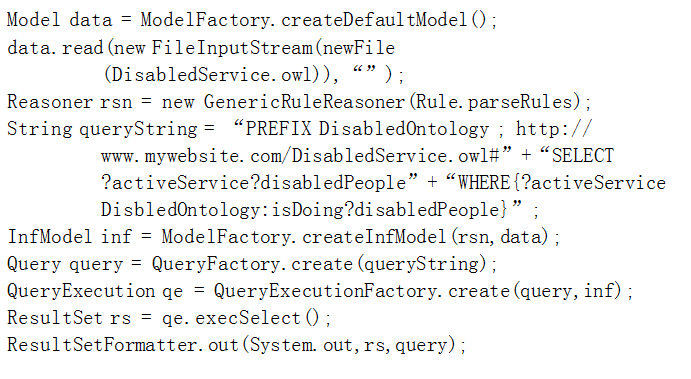
推理后输出结果：Reasoning Result：[Li,locate\_in , communityHosipital],因为小李在201骨科，骨科是社区医院的组成部分，因此系统判断小李在社区医院中。

（2）规则推理

系统借助SPARQL查询语言实现了基于规则的上下文推理。SPARQL是Jena推荐的本体查询语言，它可以用于描述规则，以及实现基于规则的推理[43]。下面是一条规则Rule1，描述了残疾人患有高血压病，且当前时间是晚上，且血压值处于异常状态，则判断应向该用户提供主动提示，提醒其服用降压药物。

Rule1：（？disabledPeople hasDisease ？hypertension）^(?time is evening)^（？bloodPressure hasStatus abnormal）->(?activeService provideTo disabledPeople)

系统通过加载规则来得到推理结果，实现代码如下：



经过推理，返回的结果如下所示：



推理的结果说明，系统会向用户发出应按时服药的提示，这与实际情况是一致的。