文章编号:1001-2486(2011)03-0099-06

# 探索性分析工具 EASim 的设计与实现*

雷永林,赵 新,张 伟,李 群
(国防科技大学 信息系统与管理学院,湖南 长沙 410073)

**摘 要:**探索性分析是一种研究深度不确定条件下复杂高层问题的分析方法。为使其能够有效地解决实际问题,需要探索性分析工具的支持。国内目前的探索性分析工具大都是面向特定应用开发的,缺乏规范化的探索性分析模型描述方法。采用扩展影响图作为探索性分析模型的统一描述规范,自主开发了通用型探索性分析工具 EASim。介绍了 EASim 的功能结构、关键技术和软件实现,指出了 EASim 较之国外同类软件 Analytica 的技术特色。多家单位的应用实践验证了 EASim 的有效性。

**关键词:**探索性分析;扩展影响图;EASim
**中图分类号:**TP391 **文献标识码:**A

# Design and Implementation of Exploratory Analysis Software EASim

*LEI Yong-lin, ZHAO Xin, ZHANG Wei, LI Qun*
(College of Information Systems and Management, National Univ. of Defense Technology, Changsha 410073, China)

**Abstract:** Exploratory Analysis (EA) is a methodology for complex macroscopical problems under deep uncertainty conditions. Supportive tools are necessary for the successful application of EA. Current EA software is mainly designed for specific applications, and lacks generic and formal modeling capability. A general-purpose EA tool, called EASim, was proposed. EASim uses extended influence diagram as its formal modeling specification. Its functional architecture was discussed firstly. Then the analysis and solutions of several key technologies of EASim were presented. Its software structure, runtime snapshots and technological advantages over Analytica were also presented. EASim has been adopted by several military organizations, verifying its feability.

**Key words:** exploratory analysis; extended influence diagram; EASim

探索性分析是近年来美国兰德公司针对不确定条件下复杂高层问题研究提出的分析方法论[1]。其基本思想是,在多分辨率模型的基础上,通过对各种不确定性要素所产生的结果进行整体研究,在较高层次模型上进行的相关探索,在深入研究问题细节之前,获得问题在广度上的影响关系。当前,国际上对探索性分析研究较为深入的当属兰德公司科学家 Davis 等人,其采用的分析工具主要是 Analytica[2]。国内许多单位都开展了探索性分析的研究,从公开发表的文献来看,尚缺乏支持分析模型描述和不确定性处理的软件框架。现有的软件工具[3-5]主要面向特定应用,缺乏灵活的分析模型描述方法和描述语言规范,适用范围有限。为支持探索性分析的研究与应用,我们结合相关应用单位的实际需求,借鉴 Analytica 的建模规范和应用模式,自主开发了通用型探索性分析工具 EASim。下面首先讨论 EASim 的设计思路与功能结构,然后讨论其中关键技术的设计方案,最后给出软件的实现及其较之 Analytica 的技术特色。

## 1 EASim 的功能结构

探索性分析的一般过程为:1)对影响问题的关键因素进行提取,勾勒各因素之间的因果关系,建立定性的问题分析树;2)将影响因素细分为能力因素、环境因素、策略因素、中间因素、目标因素等类型,以数组或概率分布的形式给出各因素的定量描述,根据机理或统计方法建立因素之间定量的影响关系模型;3)对生成的想定空间进行缩减,剔除无需探索的空间点;4)对剩余的想定集进行运行解算,一次性得到想定集内所有想定的运行结果;5)对运行结果进行探索性分析,寻求鲁棒解。

探索性分析工具应为上述过程所需的功能提

* 收稿日期:2010-10-25
基金项目:国家自然科学基金资助项目(60974073,60974074);国家省部重点基金资助项目(9140A04010106KG01)
作者简介:雷永林(1978—),男,讲师,博士。

供有效支持。参照 Analytica 的成功经验,我们设计的 EASim 的功能结构与运作过程如图 1 所示。

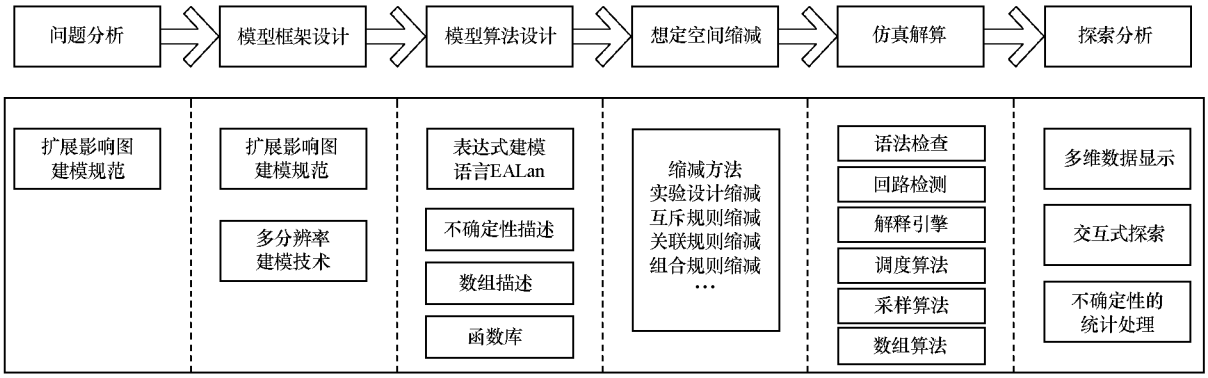


图 1 EASim 功能结构

Fig.1 The function architecture of EASim

整个结构按照系统运行过程分为问题分析、模型框架设计、模型算法设计、实验设计、仿真解算、探索分析等六个阶段,每个阶段 EASim 系统提供或实现的关键技术对应地列在下方方框内。

(1)问题分析阶段。本阶段对所研究的问题进行分析分解,确定问题边界以及要分析的主要指标,识别主要的影响因素,基于 EASim 提供的扩展影响图建模规范进行描述。

(2)模型框架设计阶段。扩展影响图建模方法是典型的基于变量的模型描述方法,问题分析阶段识别的主要指标和主要因素可以自动成为分析模型的基本模型框架。本阶段一般是根据对问题的深入分析,基于 EASim 扩展影响图建模规范提供的层次化模块组织功能对基本模型框架进行细化与层次化组织。根据分析的需要也可以采用多分辨率技术对模型框架进行组织。

(3)模型算法设计阶段。本阶段是对模型框架设计阶段形成的模型框架中的各个节点的算法进行详细设计,设计主要使用 EASim 系统提供的表达式建模语言 EALan。为支持常见不确定性的描述,EALan 提供了初步的不确定性描述功能,支持多种连续和离散分布随机变量的生成。为支持动态反馈仿真分析,EALan 提供 Dynamic 语句,可对动态变量及变量之间的动态相互影响进行仿真描述。通过[]运算符,支持对变量早期时刻值的访问。为支持探索空间的生成,EALan 支持数组型表达式的表示,可以通过索引型节点设置普通数组、等间隔数组、字符串列表等的描述。此外 EALan 还支持常见数学函数的直接使用,支持基于 EALan 自定义函数。

(4)想定空间缩减阶段。想定空间缩减并不是必须的。EASim 默认情况下会进行各数组形式变量的全面组合运算。考虑到总运行次数还要乘以随机采样次数,许多情况下全面组合运行会导致巨大的计算需求。为了保证实验覆盖性,可以通过均匀设计、正交设计等实验设计方法对想定空间数进行筛减处理,也可以根据具体问题的特点,基于规则进行互斥缩减、关联缩减、组合缩减等。

(5)仿真解算阶段。仿真解算是对扩展影响图建模规范与 EALan 语言的计算机实现,是对 EASim 探索性仿真分析模型的求解。仿真解算首先要进行语法检查,确保符合 EALan 语言语法规范。其次要进行回路检测,确保相同时刻的多个动态变量之间不会相互访问。通过语法检查和回路检测之后的模型将被动态翻译为 Python 语句,这是解释引擎的任务。调度算法负责仿真计算的因果关系,即确保输入变量先计算,确保首先获取到变量较早时刻值后再进行当前变量的求解。对于随机模型,应先在 EASim 中设置采样方法和采样次数,采样方法支持蒙特卡洛采样、随机拉丁方采样、中值拉丁方采样等几种。对于表达式为数组的节点,多维数组算法负责进行相应的调度与算法处理。

(6)探索分析阶段。对于多维随机动态分析模型,EASim 提供的探索分析手段对于分析人员非常重要。EASim 提供对多维数据的多维曲线可视化显示功能,并且可以方便地调整观测维度交互式探索。通过内置的统计运算算法,可以方便地给出随机因素的中值、均值、概率密度、累积概率密度、置信区间半长等数值或曲线显示。

## 2 EASim 的关键技术

EASim 上述功能架构中主要涉及三方面的关

键技术,即扩展影响图建模规范、EALan 建模语言以及整个模型的运行解算,下面分别给出其设计方案。

## 2.1　扩展影响图模型规范

影响图(Influence Diagram)是在决策分析等理论中建立的反映决策变量与输入变量和输出变量之间影响关系的有向图[6]。经典影响图的节点包括决策节点、机会节点和效用节点三类。在此基础上,EASim 结合探索性分析的需要进行了扩展,支持以下九类节点和两个内置变量。

(1)决策变量节点:由决策人员直接控制的变量,不依赖于其它变量节点,可以是常数、常数列表或字符列表。一般作为其它变量的输入变量。

(2)随机变量节点:由随机分布函数定义的变量,或定义中包含随机分布函数的变量。一般作为其它变量的输入变量,也可以接受其它变量为输入。

(3)一般变量节点:具体类型尚未确定,或者决策人员不能直接控制而又不是随机变量。如果其它类型的节点更合适,应改为其它类型。

(4)常量节点:该变量的值是固定不变的,没有输入变量。

(5)索引变量节点:一般定义为数组的一个维度,在结果显示时作为 X 轴或多维曲线的键。索引变量可通过名字访问,仿真解算模块会自动遍历索引变量的每个值,给出相应的计算结果值。

(6)目标变量节点:也称作效用节点,一般用于描述模型的总体指标变量,一个模型一般有一个目标变量节点。

(7)函数节点:EASim 内部函数不能满足要求时建模人员可以通过本节点定义自己的函数。

(8)模块节点:包含一组变量和其它子模块,每个模块对应一个模型图。模块节点是支持扩展影响图层次化描述的关键。

(9)代理节点:可以为上述 1 – 7 型节点创建代理节点,代理节点没有自己的定义,直接引用所代理的节点,主要用于在不同的模块中引用其它模块中的变量,也是扩展影响图层次化描述的机制。

(10)仿真时间变量:是系统内置的索引型变量节点,不能修改其变量类型和名字,可以修改其表达式定义。时间变量不在模型图上显示,可以通过 Time 名字访问,也可以通过关键词 Me 隐式访问,通过 Varname[Time – k]可访问变量 Varname 的较早时刻值。

(11)采样次数变量:是系统内置的索引型变量节点,不能修改其变量类型和名字,可以修改其表达式定义。不在模型图上显示,可以通过 Run 名字访问。

## 2.2　EASim 的建模语言

(1)基于 Python 的建模语言架构

EASim 建模语言(称作 EALan)用于支持上述各型变量节点的算法表达式定义。EALan 的语法基于巴卡斯范式(BNF)定义,执行语义实现则映射到 Python 脚本语言。Python 是一种强大的脚本编程语言[7],自上世纪 90 年代中期出现以来,已经广泛用于许多领域,特别是在科学计算研究领域应用最广的脚本编程语言,在全世界拥有广泛的用户群,Python 语言模块所实现的众多函数库可以大大简化 EALan 的实现。此外,Python 提供了 C/C++ 语言的扩展和嵌入接口,可以将 Python 引擎嵌入到应用程序中,也可以根据需要通过 C/C++ 扩展 Python 的功能。为此我们选择 Python 语言作为 EALan 的中间表示语言,通过将 EALan 语言翻译为 Python 语句,使用 Python 引擎的嵌入功能,将 Python 引擎嵌入到 EASim 系统中,实现对 EALan 语言的解释执行;并基于 Python 的 C/C++ 扩展接口来实现任意外部动态链接库(DLL)中计算函数的配置式无缝集成。EALan 及其解释引擎的总体架构如图 2 所示。

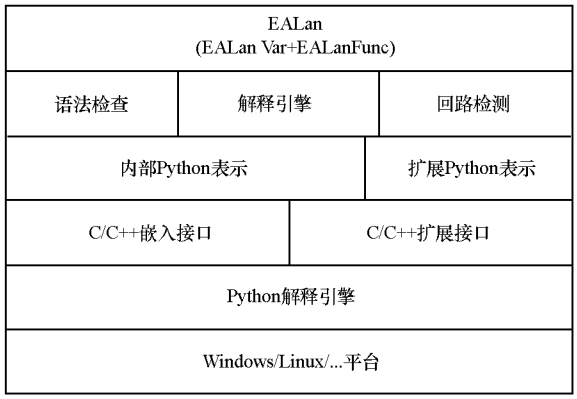


图 2　EALan 及其解释引擎的总体架构
Fig.2　The architecture of EALan and its interpretation engine

(2)EALan 的语法设计

EALan 包括 EALanVar 和 EALanFunc 两部分,前者用于描述各型变量节点的表达式定义,后者支持自定义函数的表达式定义。二者均包括词法规范和语法规范两部分,篇幅限制,不再给出。EALan 的主要语法结构包括下面 6 个方面:1)普通表达式:支持变量间的简单算术运算以及对运算结果的逻辑判断。2)内置数学函数:支持各种

常用数学函数的使用。3)概率分布函数:支持各种概率分布函数的使用。4)自定义函数:支持自定义函数的定义和使用,且自定义函数的定义中可以使用其它自定义函数和内置数学函数。5)数组运算:支持决策变量以数组形式表示各种取值水平。6)动态反馈:支持对变量较早时刻值的访问,提供动态反馈能力。

(3)EALan 的语义实现

EALan 的语义实现就是将 EALan 语句翻译为对应的 Python 语言表示,具体采用 Lex&Yacc 工具[8]进行语法编辑和语义生成。上述 6 个语法结构中前 3 个都可以相对直接地映射到 Python 语言(Python 本身内置了大量的数学函数和概率分布函数)。后 3 个则需要进行针对性的翻译设计。

自定义函数的语义实现可以类似变量进行翻译,区别在于,每个自定义函数中只能访问该函数定义的参数、内置的数学函数以及其它自定义函数,不能访问模型中的变量,且要确保和所使用的自定义函数之间不存在循环依赖。数组运算使得 EASim 可以支持大量探索空间的表示与计算。每个数组由一个或多个索引型变量所定义,多个数组间的运算也是基于数组所使用的索引变量进行的。多个数组运算结果的索引变量是参与运算的数组的索引的并集。形式地表示如下:设数组 $A$ 的索引集是 $I_a = \{U_1, U_2, \cdots, U_n\}, n >= 0$,另一个数组 $B$ 的索引集是 $I_b = \{V_1, V_2, \cdots, V_m\}, m >= 0$,则 $A$ 和 $B$ 进行运算后的结果 $C$ 也是一个数组,其索引集为 $I_c = I_a \cup I_b$。动态反馈的功能则基于临时变量来实现,即在翻译时自动为每个访问早先时刻值的变量动态定义一个临时变量,在此基础上需要提供临时变量的动态更新算法,并确保相关动态变量间不存在同时循环依赖,即确保无不带延迟的回路。

## 2.3 EASim 的运行解算

运行解算可以对所有变量进行,一般情况仅针对目标变量进行。运行解算的基本算法流程如图 3 所示。

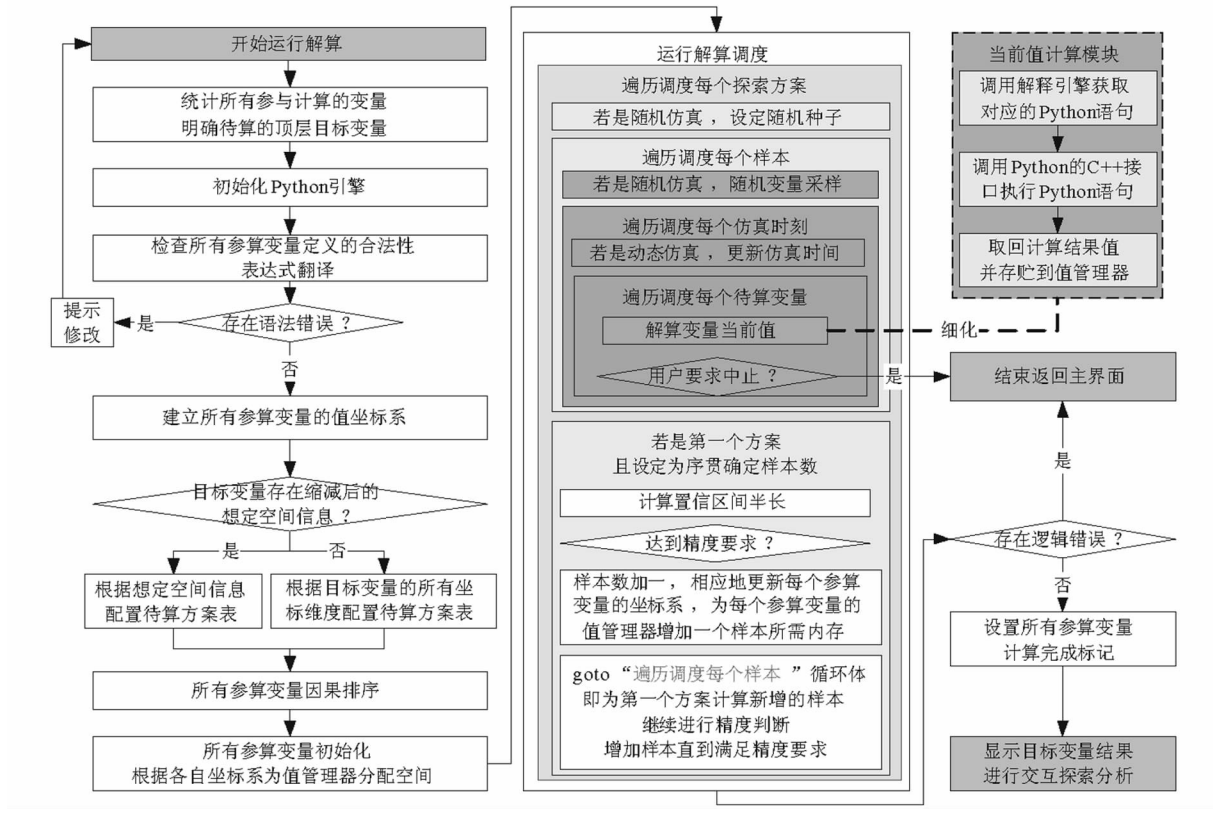


图 3 EASim 核心运行解算算法流程

Fig.3 The flowchart of EASim's kernal execution and scheduling algorithm

运行解算过程包括初始化、运行解算调度和结果显示三大部分。初始化阶段的工作主要是明确模型中的参算变量,检查参算变量定义的合法性(语法检查、回路检测),将 EALan 表达式翻译为 Python 语句,如果存在语法错误,提示用户修改。根据数组运算算法为每个变量建立存贮结果值的值管理器,以及影响变量之间值的对应关系(即值坐标系)。如果进行了想定空间缩减,则需

要按照缩减后的想定集配置待算方案表,否则根据目标变量的所有坐标维度进行全面组合配置。之后对所有参算变量按照因果关系(自变量在前、因变量在后)进行排序,后面调度时将按顺序计算每个变量的值。然后按照顺序对所有参算变量进行初始化,并为各自的值管理器分配存贮空间。运行解算调度阶段是一个四层循环体。最外层是调度每个想定(或探索方案),次外层是进行针对随机变量的采样调度,第三层是进行仿真时间调度,最内层是调度每个待算变量。每个待算变量计算时会调用当前值计算模块,该模块负责根据当前变量的当前 Python 表达式调用 Python 引擎,并从 Python 引擎中取出结果值存储在值管理器中。如果设定要进行序贯采样,则依据第一个想定的解算精度确定序贯采样的总次数,后续方案按照这个采样次数进行统一调度。正确地完成所有想定的计算后进入结果显示阶段,将调用交互探索模块对所有计算结果进行多维数据图表展现和探索性分析。

## 3　EASim 的软件实现

### 3.1　EASim 软件结构

在上述关键技术的基础上,针对功能结构设计 EASim 的软件结构如图 4 所示,分为问题分析模块、扩展影响图建模模块、想定空间缩减模块、仿真调度与计算模块、交互式探索分析模块等五大部分。问题分析模块支持分析人员定性地建立探索问题的层次分解树,并自动导出初始的扩展影响图模型。扩展影响图建模模块则载入该初始扩展影响图模型,支持进一步的定量算法定义,也可以进行对问题分析树进行层次化的多分辨率细化。在建立好的扩展影响图分析模型的基础上,可以调用想定空间缩减子系统进行想定空间缩减,扩展影响图建模模块将向空间缩减模块提供基本想定信息,后者在此基础上进行缩减,缩减后形成的实验方案传递给仿真调度与计算模块,后者按照实验方案信息进行仿真调度与模型解算,计算得到的数据可以保存到硬盘上。在计算数据的基础上可以调用交互式探索分析模块进行交互式分析探索。

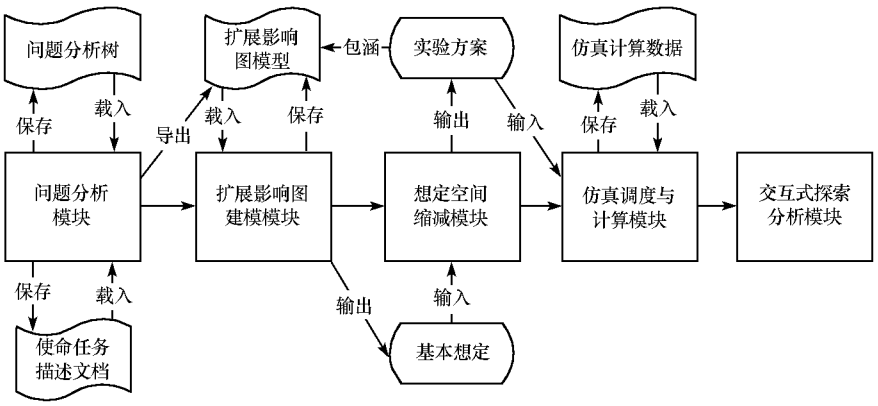


图 4　EASim 的软件结构
Fig.4　The software architecture of EASim

### 3.2　EASim 的软件实现

EASim 软件基于 C + + 语言实现。图 5 ~ 图 7 是 EASim 应用实例的截图。图 5 是显示了层次化扩展影响图模型框架,该图是某个模块的内部结构,其中的圆角矩形是其子模块,直角矩形是一般变量节点,六边形是目标变量节点,箭头表示了节点之间的影响关系。图 7 是某变量节点的表达式定义视图,图上方左侧显示了变量的类型和名字等基本信息,右侧显示的是影响当前变量的数组型变量(称作维度)列表,图中间是变量的表达式定义,图下方是影响当前变量的输入变量列表和受当前变量直接影响的输出变量列表,可以在列

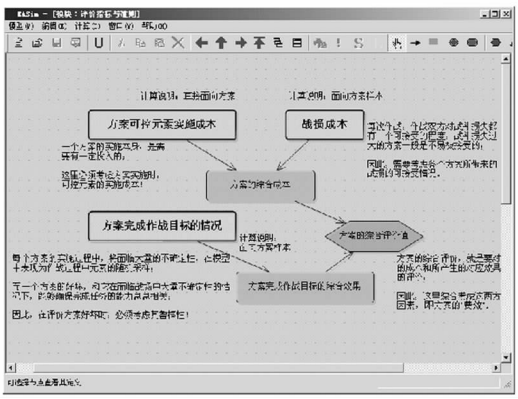


图 5　层次化扩展影响图模型
Fig.5　The snapshot of EASim' modeling view

表中进行方便的变量导航。图 6 是运行结果的显示及交互探索视图,图上方左侧是显示的数据类型及显示形式(曲线或表格),右侧是当前显示的方案信息,可以在此进行交互探索,图下方显示的当前方案下计算结果的多维数据曲线显示。

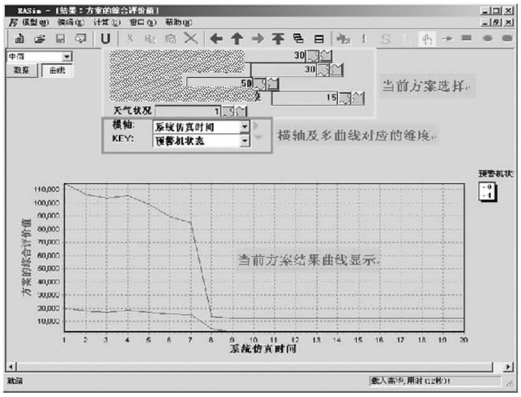


图 6 结果探索视图

Fig.6 The snapshot of EASim explorative analysis view

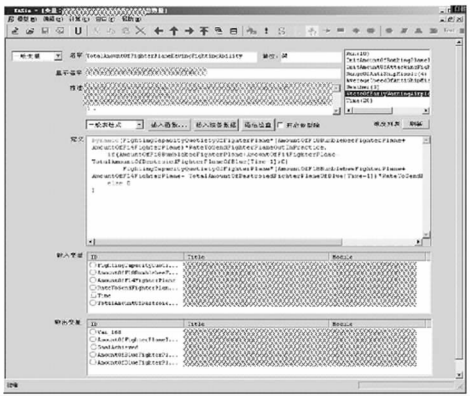

图 7 变量节点定义视图

Fig.7 The snapshot of EASim variable expression definition view

## 4 结 论

探索性分析方法离不开先进的探索性分析工具的支持。就目前的情况看,EASim 是国内第一款同时具备问题分析、探索性分析建模、想定空间缩减、仿真运行解算与交互式探索分析功能的通用型探索性分析工具。尽管 EASim 是在参照 Analytica 的基础上自主研发而成的,较之 Analytica 仍然有自身的技术特色:1)支持想定空间缩减;2)支持序贯采样技术;3)采用文件映射技术支持任意规模想定运行数据的快速存贮和检索;4)支持任意外部动态链接库(DLL)中用户已开发的计算函数的配置式无缝集成;5)支持中文。前三点对于缓解探索性分析巨大的想定空间所需的巨大计算量和巨大存贮空间非常重要;后两点对于国内人员的实际使用非常重要。EASim 在实践中已经被相关单位用于支持武器装备体系论证中的探索性分析研究。

## 参 考 文 献:

[1] Davis P K. Exploratory Analysis Enabled by Multiresolution, Multiperspective Modeling[C]//Proceedings of the 2000 Winter Simulation Conference. USA:IEEE, 2000:299 – 301.

[2] Lumina Decision Systems, Inc. Analytica 3.0 User Guide, 2003.

[3] 胡剑文,吴曦,杨镜宇,等.探索性分析框架 EAF 的设计[J]. 系统仿真学报,2005, 17(11): 2793 – 2795.

[4] 臧垒,蒋晓原,王钰,等. 基于计算实验的探索性分析框架研究[J]. 系统仿真学报,2008, 20(12):: 3077 – 3081.

[5] 柯加山, 江敬灼, 许仁杰,等. 联合作战体系对抗效能评估探索性分析框架[J].军事运筹与系统工程,2005,19(4): 58 – 61.

[6] Diehl M, Haimes Y Y. Influence Diagrams with Multiple Objectives and Tradeoff Analysis[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2004,34(3).

[7] Deitel H M, Deitel P J, Wiedermann B A. Python How To Program [M]. Prentice-Hall, Inc. 2002.

[8] MKS LEX & YACC, http://www.mkssoftware.com/products/ly, 2007.