

基于模型驱动的海量数据判读系统研究与实践

王毅, 吴晓蕊, 封慧英, 刘萍

(空间物理重点实验室, 北京 100076)

摘要: 我国“十三五”期间在运载火箭、武器型号进入高密度发射期, 在研型号进入研制关键期。高密度发射期间, 与以往发射任务相比, 发射任务繁重, 测试及发射周期明显缩短; 在测试及发射过程中, 测控数据判读是重要环节, 直接关系到飞行器各系统功能和性能判定, 整体性能评估; 传统模式下, 靶场测试、发射期间需要大量的测试判读人员现场开展数据判读分析, 耗费大量人力成本; 为解决上述问题, 提出了基于模型驱动的自动判读方法, 同时文中设计了面向特定领域的数据判读建模语言 AIPScript (AutoIP Script), 实现了姿控、制导、动力、电气总体及测量各系统或专业的判据描述、数学建模及自动判读, 并自动生成判读分析报告; 实际应用表明, 该方法的提出不仅节省了人力资源而且大大地提高了判读效率和准确度。

关键词: 模型驱动; AIPScript; 数据判读

Research and Practice of Massive Data Interpretation System Based on Model Driven

Wang Yi, Wu Xiaorui, Feng Huiying, Liu Ping

(Key Laboratory of Science and Technology on Space Physics, Beijing 100076, China)

Abstract: During the “13th Five-Year” period, the launch vehicle and weapon models entered the high density launch phase and entered the critical period of development. During the high-density launch, compared with the previous launch mission, the launch mission is heavy, and the test and launch cycle is obviously shortened. In the process of testing and launching, the interpretation of TT & C data is an important part, which is directly related to the function and performance evaluation of aircraft systems, and the overall performance evaluation. In the traditional mode, during the shooting range test and launching, a large number of test interpreters need to carry out data interpretation and analysis, which costs a lot of manpower cost. In order to solve the above problems, this paper proposes an automatic interpretation method based on model driven. At the same time, a data interpretation modeling language AIPScript (AutoIP Script) for specific domain is designed in this paper, which realizes the judgement of the attitude control, guidance, power, electric population and the system or specialty of measurement, mathematical modeling and automatic interpretation. And the interpretation analysis report is automatically generated. Practical application shows that the proposed method not only saves human resources, but also greatly improves the efficiency and accuracy of interpretation.

Keywords: model driven; AIPScript; data interpretation

0 引言

现在新型号普遍采用高码率测量体制, 参数多、数据量大, 人工数据判读、编写判读报告的工作量大、判读时间长、误判与漏判风险增加。

尤其是姿控与制导专业, 由于负责判读的参数多, 判读工作量更大, 该问题更加明显与突出, 已成为影响数据判读效率的瓶颈。传统的判读方式为人工判读, 判据多靠自然语言进行描述, 如何将判据描述精确转换为计算机语言描述, 是实现自动判读的一个关键技术, 只有使用计算

机描述语言对判据进行了定义, 才有可能实现全系统参数的自动判读。针对飞行器测试数据量大, 处理方法复杂的问题, 开展了海量测试数据处理方法、数据自动判读算法研究工作, 提升了型号自动化测试水平。本文提出的基于模型驱动设计的数据判读系统用于对各类试验数据的管理、判读与分析。基于网络实现测试数据的查询、浏览、分析、判读和报告生成等功能。软件支持自动判读与人工判读两种模式, 自动判读模式下, 通过预先设置好的判据实现全部参数的自动判读与自动报告生成, 以提高判读效率, 减少或避免漏判及误判。

1 建模语法设计

1.1 标识符设计

在建模语言中, 用来标识对象名称 (包括变量、函数)

收稿日期: 2018-05-22; 修回日期: 2018-06-20。

作者简介: 王毅(1984-), 男, 辽宁人, 工学硕士, 高级工程师, 主要从事软件系统设计、数据挖掘技术等方向的研究。

的有效字符序列称为标识符。在 AIPScript 建模语言中, 标识符就是一个对象的名字^[1]。

规定标识符由字母、数字和下划线 3 种字符组成, 且第一个字符必须为字母或下划线。在 AIPScript 中, 有 2 类标识符:

1) 关键字。AIPScript 语言中规定了 32 个关键字, 如 Diff, CM, CP, Statistics 等, 它们有特定的用途和含义, 不能用作变量名。

2) 系统预定义的标识符。如系统提供的库函数的名称 sin, cos 等。AIPScript 允许它们另作它用, 但这就会使它们失去原有作用, 容易产生混乱, 建议不要把它们用来作变量名。

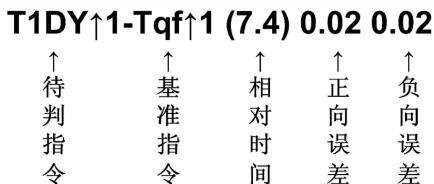
1.2 判据语法设计

判据语法设计分为两类, 分别为缓变判据语法和指令判据语法。

其中缓变判据语法共包括待判参数、起始时间、终止时间、处理公式、判据公式、正向误差及负向误差等 7 部分, 自动判读时, 将判据公式的计算结果与处理公式的计算结果进行比对。处理公式部分可以省略, 如下图所示, 当处理公式省略时, 判据公式的计算结果直接与待判参数的值进行比对。



指令判据基本语法共包括待判指令、基准指令、相对时间、正向误差及负向误差等共 5 部分, 表示待判指令与基准指令的差值应该等于相对时间的值, 误差在正负向误差范围之内。其中基准指令部分可以省略, 当基准指令省略时, 直接将待判指令的时间值与相对时间值进行比较。



1.3 表达式设计

在 AIPScript 建模语言中, 表达式设计是构成语句的基本单元, 时段、处理公式、判据公式、误差等均用到表达式。

1) 指令动作时间表达式设计: 指令动作时间表达式用于计算指令参数发生某次动作时的时间值, 其一般格式为“参数代号 开关符号 发生次序”, 其中, 参数代号为指令的代号, 开关符号表示指令的开关动作, “#”表示开, “!”

表示关, 发生次序为正整数, 表示指令动作的序号, 即指令的第几次动作。如下表所示。

表 1 指令动作时间表达式举例

参数代号	时间/s	动作	表达式
Tqf	1.0	开	Tqf# 1
	5.5	关	Tqf! 2
	10.0	开	Tqf# 3
	12.0	关	Tqf! 4

2) 差分表达式设计: 差分表达式用于计算输入表达式当前点与前某个点的差值, 其一般形式为“diff (输入表达式, 补偿量表达式, 补偿方法标识, 间隔点数)”, 其中补偿量表达式表示一个数值量, 当前前点值比前一点值小时, 根据该补偿量对差值进行补偿; 补偿方法标识用于区分不同的补偿方法, 包括“less”, “equal”, “min” 3 种方法。

1.4 判据分类设计

在 AIPScript 建模语言中, 利用标识符、判据语法和表达式, 组合成 AIPScript 建模语言中的判据语法。其中本文提出并设计了包括固定上下判据设计、参考参数判据设计、多参数运算判据设计、多时段判据设计、指令型判据设计、差分运算判据设计、丢帧分析判据设计、关键点类型判据设计、变化点类型判据设计、统计值判据设计、峰峰值判据设计和一致性判据设计, 共 12 类判据。

1) 固定上下限判据: 固定值算法的判据为一个常量, 一般用于电压、温度等物理量的判读。下面的判据给出了温度参数 W04bglq 的理论值为 293, 上下限分别为 20 的判据描述。

```
P("W04bglq");{-600,-200} {293} 20 20;
```

2) 参考参数判据: 参考参数算法的判据为另一个参数, 多用于表征同一物理意义的不同参数的比对。下例给出了模拟量采集与数字量采集的同一数据的判据描述。参数 01T01ISA1FK1_B1 的判据为参数 01R01ISA1JD_B1。

```
P("01T01ISA1FK1_B1");
{"Tqf3" ↑ 1, "ZCT-1_B1" ↑ 1}
{P("01R01ISA1JD_B1")} 0.1 0.1;
```

3) 多参数运算判据: 多参数运算判据, 多参数运算判据将多个参数进行数学运算后的结果作为待判参数的理论值。

```
P("21R01S_Ny_B1");
{"Tqf3" ↑ 1, "ZCT-1_B1" ↑ 1}
{(P("21R01S_Ny1_B1") + P("21R01S_Ny2_B1") + P("21R01S_Nz3_B1"))/3} 0.1 0.1;
```

4) 多时段判据: 多时段判据, 多时段判读语句的一般形式为:

```
待判参数: {起始时间 1, 结束时间 1} {处理公式 1} {判据公式 1} 正误差 负误差
| {起始时间 2, 结束时间 2} {处理公式 2} {判据公式 2} 正误差
```

负误差

| ...

| {起始时间 n, 结束时间 n} {处理公式 n} {判据公式 n} 正误差

负误差;

各时段内, 通过比较处理公式与判据公式的计算结果进行判断, 处理公式部分可省略, 省略时, 直接比较待判参数值与判据公式的计算结果。起始时间与结束时间可省略, 起始时间省略时, 取待判参数第一个点对应的的时间作为起始时间, 结束时间省略时, 取待判参数最后一个点对应的的时间作为结束时间。如:

```
P("U40z"): {, Tayz 1 - 0.1} {0} 0.2 0.1 | {Tayz 1 + 0.1,} {1.96} 0.2 0.2;
```

5) 指令型判据: 指令判据设计的一般形式为“待判指令动作时间表达式 基准指令动作时间表达式 理论时间表达式 正误差 负误差 时间标识符”, 其中基准指令动作时间表达式可省略, 省略时, 用待判指令动作时间与理论时间比较, 标识符包括 yctime 与 kztime 两个, 分别表示取遥测时间与取控制时间, 默认取遥测时间。如:

```
"DGL1" 1 {1} 0.02 0.02 kztime;
"DGL1"! 2 "DGL1" 1 {0.2} 0.02 0.02 kztime;
"DGL2" 1 "DGL1" 1 {0} 0.02 0.02 kztime;
"DGL2"! 2 "DGL1" 1 {0.2} 0.02 0.02 kztime;
"DGL3" 1 "DGL1" 1 {0} 0.02 0.02 kztime;
"DGL3"! 2 "DGL1" 1 {0.2} 0.02 0.02 kztime;
```

6) 差分运算判据: 差分运算判据设计, 实现计算待判参数的差分, 将计算结果与给定的值进行比较, 下例给出参数 21R01G_Ngx1_B1 的差分判据描述。Diff 运算符标识差分运算, Diff 函数中第一个参数标识计算差分的参数, 第二个参数标识差分计算的点间隔, 50 表示当前点与当前点的前 50 个点作差值, 第三个参数标识参数的溢出补偿; 函数 T 标识取参数当前点或相对当前点指定位置点的时间; 0.002674 标识差分计算的的理论值。

```
P("21R01G_Ngx1_B1"):
{,}
{
(diff("21R01G_Ngx1_B1", 50, 1.6))/
(T("21R01G_Ngx1_B1"[0]) - T("21R01G_Ngx1_B1"[-50]))
}
{0.002674}
0.1 0.1;
```

7) 丢帧分析判据: 丢帧分析判据是差分计算的一个特例, 特定用于帧计数参数, 判断每个帧计数和前一个帧计数之间的差值是否为 1, 以此确认是否有丢帧情况, 其中待判参数一般配置为帧计数, 比如 frameCount, 判读时间起始一般不设, 默认为全时间段判读, Diff (" frameCount ", 1, 65536) 表示每一个数据点都和前一个数据点直接做差,

如果差值小于 0 则加上 65536 后再进行计算, 理论值均为 1, 表示不丢帧, 如果差值大于 1, 则表示有丢帧情况发生。

```
P("frameCount" ):
{,}
{(diff("frameCount", 1, 65536))}
{1} 0.1 0.1;
```

8) 关键点类型判据: 关键点类型判据设计用于对待判参数的指定时间点的值进行判读的判据描述。下例给出参数 21R01G_DeltaJ_B1 的关键点判据, 在判读算法部分, 给出了 3 个 (时间, 值) 的数对, 分别表示在 3 个关机时刻点对应的 21R01G_DeltaJ_B1 的理论值。

在关键点类型判据中, 如果理论值公式 (即" 21R01G_T_B1") 为空, 则取参数的遥测时间为关键点的查询依据, 如果理论值公式不为空, 则以理论值公式计算结果作为关键点的查询依据。如下例中, 在获取第一个关键点值的时候, 首先计算指令" 04T01Tk1_B1" ↑1 的时间 T, 然后在参数 21R01G_T_B1 中查找值等于 T 时对应的遥测时间 Tyc, 在参数 21R01G_DeltaJ_B1 结果中查找 Tyc 时刻对应的值, 为最终结果。

```
P("21R01G_DeltaJ_B1")
{0, T_scope}
{
{"04T01Tk1_B1" ↑ 1, 10.8118}
{"14T01Tk2_B1" ↑ 1, 313.1792}
{"14T01Tk3_B1" ↑ 1, -1.2292}
}
{"21R01G_T_B1"}
1000 1000;
```

9) 变化点类型判据: 变化类型判据设计用于对待判参数的值发生变化的时刻进行判读。下例给出参数 21R01G_Cutoff_B1 的变化点判据, 在判读算法部分, 给出了 6 个 (值、时间) 的数对, 分别表示参数 21R01G_Cutoff_B1 变化为相应值对应的的时间的理论值。

在变化点类型判据中, 如果理论值公式 (即" 21R01G_T_B1") 为空, 则取参数的遥测时间为理论值, 如果理论值公式不为空, 则以理论值公式计算结果作为变化点的理论值。

```
P("21R01G_Cutoff_B1" )
{0, 800}
{
CP({0,0} {1,172.3073} {3,173.3073}
{11,183.43} {43,581.13} {2347,601.13})
}
{"21R01G_T_B1"}
0.6 0.6;
```

10) 统计值判据: 统计值判据设计如下, 第一个误差为最大值的上限, 第二个误差值为最小值的下限。

Abcd{,} {Statistics()} 3 1;
 Igfl {,} {Statistics()} 1.5 1.1;
 Igps {,} {Statistics()} 0.6 0.5;
 Ijzd {,} {Statistics()} 0.5 0.3;

11) 峰峰值判据: 峰峰值判据设计针对那些没有理论值, 但是有明确的上下限类型的缓变参数, min_max 为关键字, {P("U19s1")} 为处理公式, 一般可以省略。参数在给定的时间范围内最小值和最大值由 min_max (-0.1, 0.1) 指定。

PC("U19s1");{,} {P("U19s1")} min_max(-0.1, 0.1);

12) 一致性判据: 一致性判据设计使用运算符 CM 进行描述, CM 中参数个数不限, 进行判读时, 用第一个参数作为参考参数, 后续各参数均与第一个参数进行比较, 判断是否在允许误差范围之内, 下例给出 3 个导航速度参数的一致性判据描述。该类型判据在判读报告中给出一致性判读结果的表格描述, 同时给出多个参数的曲线图。

"21R01G_Vz" {,} {mp(PC("21R01G_Vz_B1"), PC("21R01G_Vz_B2"), PC("21R01G_Vz_B3")) } 0.1 0.1.

2 系统设计

海量数据自动判读系统设计包括前端和后端两部分, 前端采用 B/S 和 C/S 架构的结合实现数据判读功能, 后端采用 B/S 架构设计实现数据分析、数据存储与管理 and 数据采集设计。系统总体架构如图 1 所示^[2-4]。

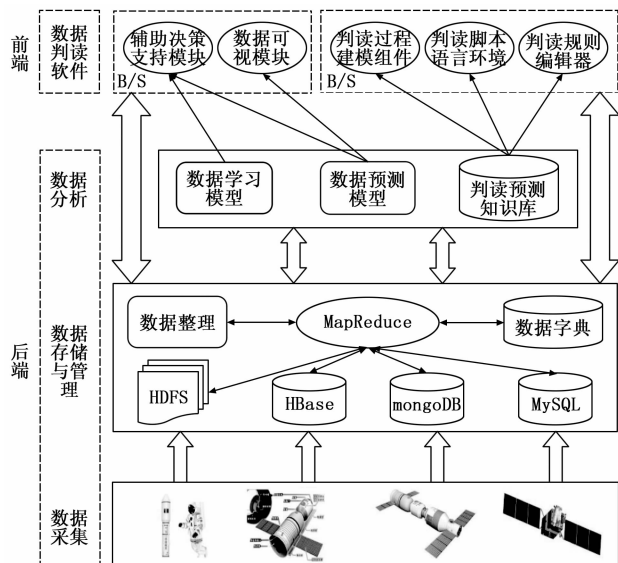


图 1 系统架构设计

数据自动判读系统的总体工作流程和原理简要表示如下: 数据判读服务器启动后, 将知识库备份中的规则和事实加载到位于内存区的黑板中。数据判读服务器通过网络, 从外部接口接收到的数据后, 首先对数据进行解析, 通过数据驱动推理引擎工作。推理引擎得到判读的配置信息, 包括判读所采取的业务类型 (参数判读、指令判读、事件判读、指令监视和事件监视等) 及其他配置信息。数据判

读服务器上的知识库副本从数据库服务器上的知识库得到刷新。此时推理机采取特定的推理策略、方法和推理算法, 调用黑板区加载的事实, 通过黑板区加载的知识库中的规则实现推理过程, 得出监视判读的结果。在推理过程中, 调度引擎模块实现对规则的动态调度。推理机将判读结果发送到解释器, 由解释器对判读结论提出解释, 说明结论得出的理由。之后解释器将判读结论及相关解释数据进行显示。

数据自动判读系统的基本工作流程和原理如图 2 所示。

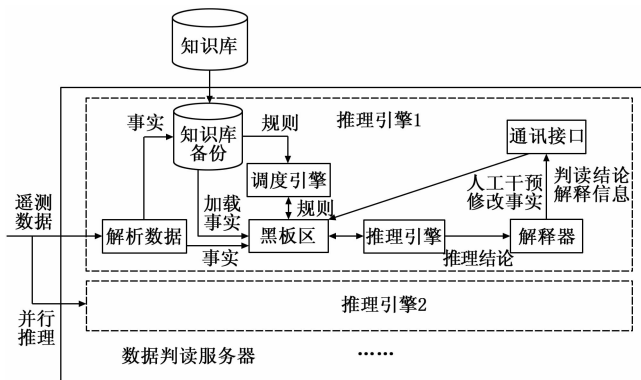


图 2 软件工作流程及原理

数据自动判读系统用于对各类试验结果数据的管理、判读与分析。主要功能包括试验选择、试验数据管理、自动判读、人工判读、模板管理、判读结果分析、报告管理、访问权限控制、设置等功能, 如图 3 所示。

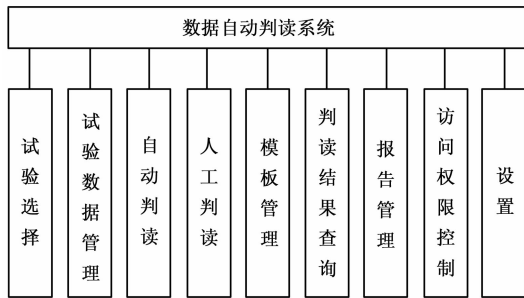


图 3 数据自动判读系统功能

3 关键技术

3.1 图形化建模技术

基于模型驱动的自动判读系统应用 AIPScript 语言驱动判断模型运行。如何将判读模型的图形化形式转化成脚本是该系统的关键技术之一。把模型中的每个判据作为一个函数 $f(x, y)$, 将各判据规则的输出转换为另一个规则的输入, 如函数 $g(x)$, $h(x)$ 的值看成规则的两个输入, 则判读模型的最终输出即为 $f(g(x), h(x))$ 的值。可视化建模工具软件提供判据和表达式元素库作为判据可视化建模的基础元素。判据的可视化建模本质上是通过拖拽相应的表达式和判据到判据编辑面板上, 并通过设置相应

的属性值来完成判据的编辑和生成的。

整个判据可视化建模工具软件主要设计了判据和表达式元素库、判据创建、判据可视化编辑、判据自动语法检查和判据保存这几大模块。图形化建模界面如图 4 所示。

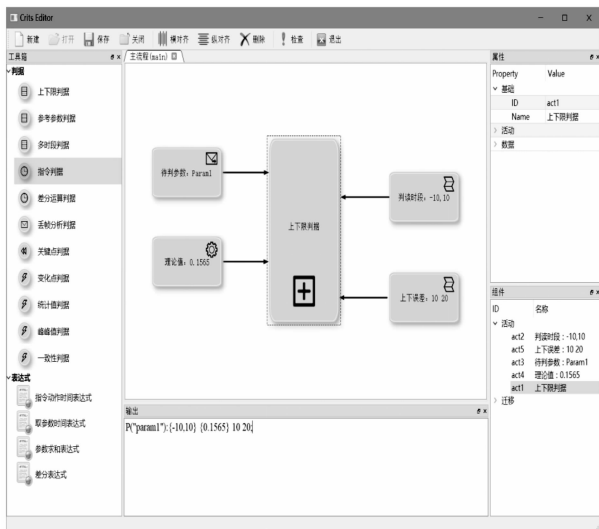


图 4 图像化建模示意图

针对峰峰值判据规则建立的判读模型后生成的 AIP-Script 脚本示例如图 5 所示。



图 5 AIPScript 脚本示意图

3.2 速/缓变参数的曲线拟合

对于在试验中需要重点关注的且变化复杂的速变或缓变参数，为了更好地实现自动或人工判读以及事后数据分析，在对该类参数进行判断规则建模时，需对该类参数进行线性或对数等曲线拟合。在进行曲线拟合时，首先需要确定理论曲线的关键点以及曲线在各关键点中的理论值，然后再根据关键点及其理论值设计合理算法并进行建模，最后得出该拟合理论曲线^[5-7]。

3.3 数据预处理

在遥测数据的接收和处理中会有野值，自动判读系统对接收到的遥测数据，首先需要进行预处理，本文采用肖

维涅准则来剔除野值，减小自动判读过程中对效率和质量的影响。

另外，对于振动数据考虑到零点漂移具有长周期趋势项，在对该类数据进行后续的判读或分析时得到的结果可能完全失真，因此利用最小二乘法消除数据趋势项，减少零漂和长周期趋势项对于后续判断和分析的影响。

4 应用效果

基于模型驱动的自动判读系统已成功应用于多个型号测试。试验前，试验人员根据型号需求对试验数据进行判据配置，判据配置包括可视化建模及脚本编辑方式，判据可根据不同专业判读需求保存判据模板，判据模板无需多次重复配置。试验中，自动判读系统根据判据对全部试验数据进行判读，对不满足判据的数据给出提示，同时提供人工判读模式对异常数据进行精细化判读，避免了试验数据的误判和漏判。试验结束后，自动判读系统生成判读报告，同时试验人员也可直接根据试验需要配置试验报告的内容，总之海量数据自动判读系统的研制大大节省了试验人员判读数据和编写报告的时间，提升了试验效率和质量，为型号研制提供了有力保障。

5 结语

在基于模型驱动的自动判读系统内，用户通过在可视化的判读规则视图中，建立用户所需要的判读规则。通过对判读参数与判读规则视图化建立，从而实现对测试数据进行自动判读，并根据条件自动输出、记录测试结果。基于模型驱动的可视化判读建模减少了数据错判与漏判，减轻了测试人员的工作量，提高了测试效率。

参考文献:

- [1] 秦 巍, 郭永富. 一种基于历史遥测数据的在轨卫星故障预警系统 [J]. 航天器工程, 2010.
- [2] 谢明明, 沈湘衡, 贺庚贤. 空间相机仿真测试数据自动判读系统的设计 [J]. 计算机测量与控制, 2010, 18 (6): 1277 - 1279.
- [3] 张 强, 郭丽丽, 马振林. 基于模型自动判读的研究与实现 [J]. 计算机技术与发展, 2014.
- [4] 吴 伟, 张 威, 潘顺良, 等. 自动判读系统在载人航天器电测中的应用 [J]. 航天器环境工程, 2011.
- [5] 赵 峰, 文博武, 孙 洁, 等. 基于卫星数据判读的一种方法 [A]. 第十六届全国测试与故障诊断技术研讨会论文集 [C]. 北京: 中国计算机自动测量与控制技术协会, 2007: 225 - 229.
- [6] 戴居峰, 袁 利, 王 跃. 基于开放式结构的航天器控制系统测试技术研究 [J]. 空间控制技术与应用, 2012, 28 (5): 34 - 37.
- [7] 吴限德, 孙兆伟, 陈卫东. 小卫星测试数据在线时间序列判读方法 [J]. 哈尔滨工程大学学报, 2012, 33 (3): 383 - 388.