

一篇文章了解STEP文件格式

原创 邓子平 [多物理场仿真技术](#)



STEP文件作为三维CAD通用文本中间交换格式，在几何数据交互中起着重要作用。但是STEP作为一种标准不仅仅只是支持几何数据信息，其中还有大量属性，行业以及扩展标准，本文从以下几个方面详细介绍STEP相关内容：

- 1. STEP简介
- 2. STEP文件格式
- 3. STEP软件应用
- 4. STEP软件研发

1.STEP简介

1983年，STEP标准由国际标准化组织(ISO)工业自动化与集成技术委员会(TC184)下属的第四分委会(简称SC4)开发。1992年，STEP标准已成为ISO国际标准(ISO10303)，国内将其制定为同名国家标准，标准号为GB/T 16656。

STEP全称是Standard for the Exchange of Product Model Data，翻译过来是“产品模型数据交换标准”，是一个比较全面的ISO标准(ISO10303)，该标准描述了如何表示和交换数字产品信息。

基于以上描述，我们可以知道：

- 1. STEP是一个标准，并不是某个特定软件或产品；
- 2. *.STEP文件是STEP标准的数据保存形式；
- 3. STEP标准由国际标准化组织ISO维护，而不是由某个公司或产品维护。
- 4. 实现读写STEP文件功能需要参照ISO发布的文档自行实现。当然有些公司提供STEP文件的读写模块，在产品开发中可以作为第三方工具集成。
- 5. STEP标准是一个通用的数字化信息标准，不仅仅只是针对CAD数据格式。
- 6. STEP是一个开放扩展的标准，从诞生之初就一直在不断维护更新。

单词缩写

AP--Application Protocol

IR-- Integrated Resource

ISO10303内容按照**组件 (Part)** 来划分，每一个组件表示一种标准，每个组件有一个唯一的数据标识号，标识号从1开始，按照“Part+标识号”唯一表示一个组件 (Part)
所有的组件内容按照内容又可以划分到7大模块中：

1. 描述方法:

Part11到Part19

2.实现方法:

Part21到Part29

3.一致性测试方法和框架:

Part31到Part39

4.集成资源通用资源:

Part41到Part99

其中Part42定义了基于三维BREP结构的几何和拓扑信息

5.集成资源应用资源:

Part101到Part199

其中Part107定义了有限单元分析信息

6.应用协议:

Part201到Part1199

7.抽象测试集:

Part1021到Part2199

按照整个框架结构，7大模块又被划分成三个层次:

其中1-3为物理层

4和5为逻辑层

6和7位应用层



多物理场仿真技术

以上图片来源于网络

在AP的基础上，还会扩展出各种衍生版本。

每个Part的详细介绍可以参考ISO文档，大概介绍可以参考wiki，链接:

https://en.wikipedia.org/wiki/ISO_10303

其中一般用户比较关心的集中在[应用协议\(Application Protocol\)](#)层, 比如常见的AP203, AP204, AP214, 支持CAM加工的AP238标准。

说明: 以下蓝色字体引用ISO标准文档

比如最常见的AP203中, 定义了以下内容:

- 属于机械零件和组件的产品;
 - 与产品开发的设计阶段有关的产品定义数据和配置控制数据;
 - 通过在子装配中的使用来表示装配中零件的实例;
 - 零件的三维形状表示, 包括:
 - 几何和拓扑有界线框模型;
 - 几何有界曲面模型;
 - 具有多面、初级和高级面的拓扑有界实体模型;
 - 非曲面有界实体模型, 包括构造实体几何、曲线扫掠和其他扫掠实体、加厚面实体;
 - 具有构建历史的实体;
 - 拓扑有界流形表面和地下和非流形表面模型;
 - 拓扑有界复合模型;
 - 几何验证属性, 允许检查几何形状表示 (高级边界表示和多面边界表示实体) 的转换质量;
 - 应用于几何形状表示的几何和尺寸公差;
 - 材料及其化学物质组成;
 - 复合材料结构和形状;
 - 以属性值对为特征的目录数据;
 - 产品数据的三维呈现:
 - 在层和组中排列几何元素并分配颜色;
 - 点、曲线、曲面和截面的表示样式, 包括阴影和平铺;
 - 保存特定摄像机位置和部分的视图;
 - 应用于几何元素的文本注释和注释;
 - 表示几何和尺寸公差;
 - 技术图纸作为产品数据的二维表示
-

而在AP204中, 规定了在机械工程设计环境中[使用和交换边界表示实体模型](#)的范围和信息要求所需的集成资源的使用。包括具体以下内容

- 用于表示形状的三种 B-rep 模型:
 1. 离散面 B-rep 模型;
 2. 具有基本曲面的 B-rep 模型;
 3. 具有雕刻表面的 B-rep 模型;
- 曲线和曲面几何;
- 参数空间中定义的曲线 (pcurves) ;
- 流形拓扑;

- 产品识别信息；
- 简单的表示属性，如线型、线宽、颜色与整个 B-rep 模型的关联，或者与 B-rep 模型的几何或拓扑元素的关联；
- 保存用户定义的对象名称；
- 与几何元素相关的单位和度量；
- 零件和子组件的装配。
- 尺寸标注；
- 公差；
- 制造信息；
- 高级演示功能，例如多视图、字符字体和符号。

说明：以上蓝色字体引用ISO标准文档

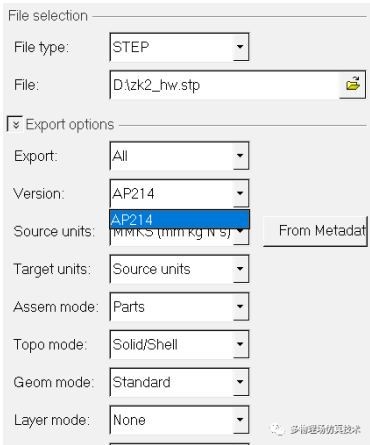
可以看出AP203，AP204定义出了较完整的基于BREP结构的三维模型数据，目前最新的版本是2020年修订的AP242版本，链接：
<https://www.iso.org/standard/66654.html>

AP242向上兼容，也就是说可以用老版本接口可以读入，但是新的属性信息无法读取。对于读写STEP文件，需要明确知道产品或组件支持的STEP标准版本信息，如果使用的读写标准版本比较老，新标准的数据是无法读出的。

目前商业软件一般对AP203，AP214，AP242导入都支持的比较好，但如果版本较老，AP242标准里有些后来新加的属性可能支持的不是很好，导出一般只支持到AP214。

开源的三维几何内核OCC到7.3版本导入只支持到AP214，AP242部分支持。

下图是Hypermesh的STEP导出标准，只支持AP214。



另外，对于各种设计仿真软件，如果几何导入功能只能支持STEP/IGS，而不能支持SAT,X_T的，基本可以确定几何内核使用的是免费OCC(OCC提供商业模块可以支持SAT和X_T)

下面的列表显示了 STEP 应用程序协议 (Application Protocol AP)，它们都建立在同一组集成资源 (Integrated Resource IR) 之上，因此它们都对相同的信息使用相同的定义。例如，AP203、AP214、AP238 和 AP242 对三维几何、装配数据和基本产品信息使用相同的定义。

- AP 201: 显式绘图
- AP 202: 关联绘图
- AP 203: 配置控制设计
- AP 203e2: 配置控制设计 (第二版)

AP 209: 复合材料和金属结构分析及相关设计

AP 210: 电气连接包装设计

AP 214: 汽车设计

AP 215: 船舶布置

AP 216: 船模

AP 218: 船舶结构

AP 219: 尺寸检测信息交换

AP 221: 过程工厂的功能数据和示意图

AP 223: 铸件

AP 224: 基于特征的流程规划

AP 225: 建筑元素

AP 227: 植物空间配置

AP 232: 技术数据打包

AP 235: 工程属性

AP 236: 家具目录和室内设计

AP 238: STEP-NC 集成 CNC

AP 239: 产品生命周期支持

AP 240: 宏观流程规划

AP 242: 基于托管模型的3D工程


AP 242e2 容差处理

STEP和IGS的关系





两者没有关系，IGS是早于STEP一种文件标准格式，但由于携带信息量少，数据不完善，转换信息容易丢失等原因，除了历史因素部分软件还保留支持外，工业界已经普遍转向STEP标准。IGS标准可以认为是STEP标准的前任。

2. STEP文件格式

我们可以在任意三维CAD软件中绘制一个长方体，然后将其导出为STEP文件格式。这里以FasCAD为例。

 **START** **GEOMETRY** **2D GEOMETRY** **3D GEOMETRY** **TOOL** **HELP**

Box	Pyramid	Truncated-Cone	Partial Cylinder	Partial Truncated-Cone	Extrude by Loop	Loft	Paraboloid Cone	Build
Cylinder	Sphere	Torus	Partial Sphere	Partial torous	Sweep		Cylindrical Pipe	PCB
Wedge	Cone		Partial Cone	Extrude	Revolve		Turbine	
3D Primitive			3D General			3D Custom		

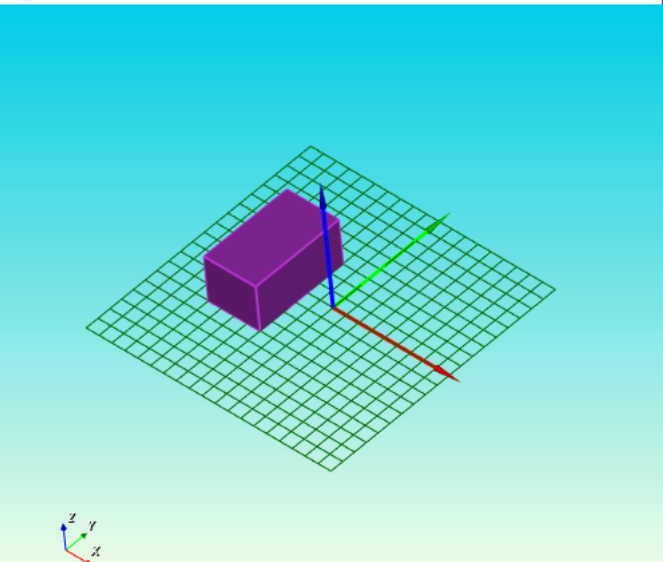
Project  **Geometry-View**   

Geometry

Box_0

Property

Property


**Message:**

2022/03/13 09:47:44 Information: Box : Input base point.

2022/03/13 09:47:45 Information: Box : Input second corner point:

2022/03/13 09:47:45 Information: Box : Input height:

2022/03/13 09:47:48 Information: Box: Done

 多物理场仿真技术

```
*demo - 副本.step - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
ISO-10303-21;
HEADER;
FILE_DESCRIPTION((' Open CASCADE Model'), '2;1');
FILE_NAME(' Open CASCADE Shape Model', '2022-03-13T08:13:21', (' Author'), (
    ' Open CASCADE'), ' Open CASCADE STEP processor 7.3', ' Open CASCADE 7.3'
    , ' Unknown');
FILE_SCHEMA((' AUTOMOTIVE_DESIGN { 1 0 10303 214 1 1 1 1 }'));
ENDSEC;
DATA;
#1 = APPLICATION_PROTOCOL_DEFINITION(' international standard',
    ' automotive_design', 2000, #2);
#2 = APPLICATION_CONTEXT(
    ' core data for automotive mechanical design processes');
#3 = SHAPE_DEFINITION_REPRESENTATION(#4, #10);
#4 = PRODUCT_DEFINITION_SHAPE('', '', #5);
#5 = PRODUCT_DEFINITION(' design', '', #6, #9);
#6 = PRODUCT_DEFINITION_FORMATION('', '', #7);
...
ENDSEC;
END-ISO-10303-21;
```

1. 文本内容以 “ISO-10303-21;” 开头，以 “END-ISO-10303-21;” 结尾，其文件结构分为两个部分：文件头和数据区。
2. 文件头内容位于标识符 “HEADER” 和最先出现的标识符 “ENDSEC” 之间，主要给出关于文件本身的基本信息，包括文件描述、文件名、作者等；

例子中FILE_SCHEMA里显示了STEP标准版本信息

{1 0 10303 214 1 1 1 1}

由上述信息可以看出，导出STEP接口使用的是OpenCascade7.3版本，支持的STEP标准为AP214。

3. 数据区以关键字 “DATA” 作为开始标识，同样以 “ENDSEC” 作为结束标识。

数据区每行以 “#+数字标号” 开始， “=” 后接关键字，括号内接具体数据。具体数据可以索引到别的数据，也可以是具体数据。

数据区使用的是Express数据标准，有过编程经验的同学只要有STEP标准的Express关键字手册，很容易读懂数据内容。

```

#10 = ADVANCED_BREP_SHAPE_REPRESENTATION(' ', (#11, #15), #345);
#11 = AXIS2_PLACEMENT_3D(' ', #12, #13, #14);
#12 = CARTESIAN_POINT(' ', (0., 0., 0.));
#13 = DIRECTION(' ', (0., 0., 1.));
#14 = DIRECTION(' ', (1., 0., -0.));
#15 = MANIFOLD_SOLID_BREP(' ', #16);
#16 = CLOSED_SHELL(' ', (#17, #137, #237, #284, #331, #338));
#17 = ADVANCED_FACE(' ', (#18), #32, .F.);
#18 = FACE_BOUND(' ', #19, .F.);
#19 = EDGE_LOOP(' ', (#20, #55, #83, #111));
#20 = ORIENTED_EDGE(' ', *, *, #21, .F.);
#21 = EDGE_CURVE(' ', #22, #24, #26, .T.);
#22 = VERTEX_POINT(' ', #23);
#23 = CARTESIAN_POINT(' ', (-96.90328306662, -40., 0.));
#24 = VERTEX_POINT(' ', #25);
#25 = CARTESIAN_POINT(' ', (-96.90328306662, -40., 45.));

```

6个Shell

1个LOOP有4条边

每条边两个顶点

多物理场仿真技术

Express标准是STEP标准的核心内容，如果要从事解析这块工作的话，需要熟练掌握。BIM领域常用的IFC中间格式使用的也是Express标准。

3. STEP软件应用

一般的工业设计仿真软件包括前处理软件都支持STEP文件格式的导入，但由于不同软件对STEP格式的标准，单位，容差，导入导出等一系列相关属性设置不同，很容易导致在操作过程中出现数值精度误差，几何拓扑错误，信息丢失，以及性能问题。

在处理STEP文件的时候，注意以下几个原则：

1. 在建模的过程中，尽量使用原生的功能和相关规范。比如最简单的长方体，面拉伸，扫掠，放样，布尔运算等都可以生成长方体。显然面拉伸是最简单也是推荐的，因为后面几种方法对于复杂模型很容易产生数值精度问题，在导出时精度设置不够容易出现数值精度丢失。
2. 软件尽可能使用较新的版本，但不用最新。相对较新的版本对于各种规范都支持的比较好，也有更好的兼容性。
3. 对于固化流程的文件导出导入，建议使用脚本，一般的CAD软件都有几何修复和清理功能，且支持脚本。使用脚本功能可以减少文件转换过程中属性设置造成的转换误差。
4. 注意模型单位保持一直，不同单位转换经常会造成精度丢失，导致在模型检查计算，网格划分中出现错误。
5. 根据实际需要设置是否合并模型，根据大模型分治方法，原则上将模型的组件分开处理比较合适，最主要的是可以提升各种操作性能。由于各个软件处理的最终目的不同，因此需要根据实际情况处理，比如模型在前处理软件中，需要将几何实体数据默认离散成面，但在CAD软件中可能需要以实体或装配体表示。
6. 文件导入导出过程中，注意保持标准版本一致。

STEP文件后缀通常为*.stp, *.step等，其实一般来讲，后缀名除了用来识别文件格式外，并不影响文件读写。将一个step文件改成其它后缀格式，仍然可以使用STEP文件读写器正常读写。有些软件以*.stp, *.ste, *.p21做后缀，也是STEP标准。

4. STEP软件研发

对于导入STEP文件的功能，软件需要有几何数据格式定义，这样才能将STEP里的相关信息转换成软件自己的数据结构，所以通常需要一个三维几何内核，前处理软件对几何内核的要求会低一点，CAD软件则需要尽可能保留原始信息。

STEP文件的读写对于研发来说，是最基本的操作。由于STEP格式支持的标准众多，不同行业对STEP的标准可能不太一样。但是基本的三维BREP格式都是相同的，实际上只要能用string, int, double等数据类型表达的信息都可以方便的放到STEP文件中。


从0开始研发读写STEP文件功能是一项繁杂的体力脑力劳动，一般建议使用第三方工具。

数据结构优化

EXPRESS是基于schema格式，每行都是一条信息，不同信息通过索引关联起来。也就是说只有在所有信息全都读取完成后才能进行信息重建，对于大模型来讲，这个过程是非常缓慢的，容易出现性能瓶颈。主流的三维CAD以及仿真软件读取一个几百M的STEP文件都非常慢，原因就在此。

前面举的例子中，一个简单的长方形实体的几何数据就有350行！

```
#344 = ORIENTED_EDGE(' ', *, *, #310, .F.);
#345 = ( GEOMETRIC_REPRESENTATION_CONTEXT(3)
GLOBAL_UNCERTAINTY_ASSIGNED_CONTEXT((#349)) GLOBAL_UNIT_ASSIGNED_CONTEXT
((#346, #347, #348)) REPRESENTATION_CONTEXT('Context #1',
'3D Context with UNIT and UNCERTAINTY') );
#346 = ( LENGTH_UNIT() NAMED_UNIT(*) SI_UNIT(.MILLI., .METRE.) );
#347 = ( NAMED_UNIT(*) PLANE_ANGLE_UNIT() SI_UNIT($, .RADIAN.) );
#348 = ( NAMED_UNIT(*) SI_UNIT($, .STERADIAN.) SOLID_ANGLE_UNIT() );
#349 = UNCERTAINTY_MEASURE_WITH_UNIT(LENGTH_MEASURE(1.E-07), #346,
'distance_accuracy_value', 'confusion accuracy');
#350 = PRODUCT_RELATED_PRODUCT_CATEGORY('part', $, (#7));
ENDSEC;
END-ISO-10303-21;
```

 多物理场仿真技术

可以从三个方面对STEP文件进行优化：

首先是排序，即把同一个对象的数据，属性等按照组放在一起，并删除重复的数据，读完即可进行解析转换成程序数据，不需要等到文件解析完成一直占用资源，这种方法对于大部分文件都有效，实际测试结果表明性能普遍能提高20-40%，而且保持了STEP格式，具有通用性；

第二种方法是在第一种方法的基础上将STEP文件进行拆分，比如一个文件分成10个小文件，读写的时候使用多线程，在多核多线程的机器上性能提升效果非常明显；

第三种是讲STEP文件转成自己的数据格式，然后使用二进制格式保存。二进制由于采用了按字节读写方法，读写性能要远高于文本格式读写。为了节省空间，可以使用7z等算法对数据进行压缩；另外一些STEP轻量化工具其本质也是提取STEP内的部分信息，重新组织数据结构输出。

以上方法可以混合使用，不仅针对STEP文件格式，对于一般的文件存储也可以使用，尤其是超大规模模型的时候。

之前有朋友私信问，这种标准是否会被卡脖子，有没有必要自己开发相应的中间标准格式？

其实真没必要动不动就喊“卡脖子”，卡脖子技术有确切的领域和技术点，并有其鲜明的特点：

1. 技术门槛高，起步投入资金多；
2. 研发耗时费力，需要长期积累，3年起步，成熟5-10年非常正常；
3. 花钱买不到，买到只能，无法转成自己的东西；
4. 保密，专利核心技术点多。即使投入研发也不一定做得出来；
5. 在行业内具有垄断性，无可替代

类似STEP标准以上一条都沾不到。STEP的设计初衷就是中间格式，最大程度地支持标准性，开放性和扩展性。不过国内开发新的标准并不是没有意义，只要生态做好了，仍然可以在数据标准领域占有一席之地。

从现在工业软件研发角度和发展趋势看，STEP标准已经落后于时代了，其主要表现在数据臃肿，冗余度高，因为要兼容各种信息牺牲了灵活性和敏捷性，在性能，并发，效率上已经跟不上软硬件的发展，在很多场合也满足不了实际工业应用需求。但是制定新的标准又是一项漫长繁琐的工作，短期内很难出现STEP的替代标准。

参考资料

题外说明：文中部分内容引用了附录中的资料，主要是方便读者参考。之前有接到投诉说有原创文章抄袭，本公众号一不引流，二不做广告，三不拉赞助，发行业内的招聘信息免费，致力于工业软件研发技术推广和普及，以及行业内信息研判，就像一位朋友来信说的，公司研发部系统学习了公众号的内容后，整体研发水平都有了提升；还有投行直接基于公众号内容写研报。这些说法可能有点夸张，但是能帮助到国内工业软件类研发公司，笔者还是比较欣慰的。

最后，在以后文章中，对于引用的内容，尽量给出出处，网上引用的信息也会给出信息源，总之做到合规，写出更好的原创文章。

阅读： null

在看： null