Automatic Enrichment of Ontology for Engineering Design Process

Kobkaew Opasjumruskit
German Aerospace Center (DLR)
Institute of Data Science

SEMANTiCS 2019, September 9-12
Karlsruhe
Agenda

• Scenario

• Ontology for Satellite Design Process
  – Structure & Current Development

• Automatic Improvement of Ontology
  – ConTrOn – Continuously Trained Ontology

• Summary & Outlook
Concurrent Engineering Facility

Product Specification

Model Based System Engineering Tool

Introduction
Ontology
ConTrOn
Summary

Use Scenario
Challenge #1 Heterogeneous Formats

OVERVIEW

Under a grant from the Defense Production Act Title III, Ball is developing a new line of affordable, fully-domestic star trackers, CT-2020.

Domain-specific, secure solution

Using cutting-edge technology, secure systems and flight software, the CT-2020 is an assured, fully-US-based solution for the nation’s most important missions.

Low cost, performance

Bending medium and high accuracy star tracker heritage in a compact, fully-integrated package, CT-2020 offers high performance and operational flexibility at a competitive price point.

CT-2020 integrates the latest high-precision, commercial-off-the-shelf (COTS), detector technology developed in the U.S. specifically for star trackers, enabling the CT-2020’s open-defense small size and volume design.

Operational flexibility, on-orbit adaptability

Featuring operational flexibility, CT-2020 provides customers two modes of operation: fully autonomous and directed search, in which the user can select regions of interest. In autonomous mode, the tracker can perform single or multi-axis tracking, which is an important capability for vehicles with higher accuracies in directed sensor mode.

CT-2020’s rugged software features an air or rail environment simulation, allowing the tracker to simulate mission-specific integration and operations for vibration. In addition, the tracker’s software supports updates to the star catalog, satellite intensity calibration and software algorithms.

HERITAGE

For more than 40 years, Ball has delivered the highest-reliability, highest-performance star trackers available to support U.S. commercial and defense missions. CT-2020 leverages this heritage to optimize the CT-2020 for cost and performance to bring an affordable, domestic star tracker solution to the U.S. market.

Expected availability of the CT-2020 is fourth quarter of 2019.

SYSTEM COMPONENTS

[Image of CT-2020 components]

• 1-axis performance (stand-alone unit)
• Full-time, 3-axis, <1° accuracy performance with typical mission accuracy assurance
• Full performance with a 15 deg yaw angle
• <0.1° accuracy with high tolerance at 15° yaw
• high rate capability (up to 68 degrees) with reduced performance to 0.1° angular rate through slew

Ontology

ConTrOn

Summary

Use Scenario

Specifications Subject to Change Without Notice

<table>
<thead>
<tr>
<th>Feature</th>
<th>M1C</th>
<th>M2D</th>
<th>M3C</th>
<th>M4D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (2σ)</td>
<td>10 arc-sec</td>
<td>1 arc-sec</td>
<td>1 arc-sec</td>
<td>1 arc-sec</td>
</tr>
<tr>
<td>Update Rate</td>
<td>50 Hz</td>
<td>45 Hz</td>
<td>25 Hz</td>
<td>25 Hz</td>
</tr>
<tr>
<td>DCR (M1C)</td>
<td>150 nA</td>
<td>150 nA</td>
<td>150 nA</td>
<td>150 nA</td>
</tr>
<tr>
<td>CRN (M2D)</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Data Rate</td>
<td>500 kbps</td>
<td>500 kbps</td>
<td>500 kbps</td>
<td>500 kbps</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>Specifications Subject to Change Without Notice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Imaging Time (TIT)</td>
<td>100 and 300 (or option)</td>
</tr>
<tr>
<td>Event Imaging Time (EIT)</td>
<td>10°/30°/60°/90°/120°/150°/180°</td>
</tr>
</tbody>
</table>

Specifications Subject to Change Without Notice
Challenge #2 Standard description

ECSS-E-ST-60-20C – Star sensor terminology and performance specification

- Not machine-interpretable
- Require manual efforts transferring data into tools
#1 ONTOLOGY FOR SATELLITE PARTS
Ontology: Sources

- Data models developed by DLR's in-house MBSE tool
  - https://github.com/virtuosoatsatellite
- Existing product description standards
- Actual product data sheets
- Interview with system engineers and manufacturers
- Current version: https://zenodo.org/record/2616374
Ontology: Hierarchical Structure

Spacecraft parts ontologies

Core

Base

External Ontologies

Supplier

Unit of Measurement

https://schema.org/Organization.ttl

https://github.com/HajoRijgersberg/OM
Spacecraft Parts Ontology: Base

- **Primary classes**
  - Part (Product)
  - Part’s attribute
  - Type of attribute

- **Primary properties**
  - “is property of”
  - “has property”
  - “has unit”
Spacecraft Parts Ontology: Core

• Common attributes for all parts
  – Mass
  – Lifetime
  – Operating Temperature
  – Width, Height, Length

• 26 attributes
Spacecraft Parts Ontology: Star Tracker

- Specific attributes to star trackers
  - Attitude accuracy
  - Field of view
  - SNR
  - Etc.

- 36 Attributes
Product Ontology: Further Usages

- Conversion to database schema
  https://gitlab.com/dlr-dw/ontocode
- Part data exchange interface
  - Web API
- Knowledge graph
  - Information retrieval
However, as time flies

People change, products change

And ontology should evolve

Auto-Improvement of ontology

But, how ...
#2 AUTO-IMPROVEMENT OF
ONTOLOGY
Information Extraction

- Natural Language Processing
- Semantic Knowledge

Auto-Improvement of ontology
Existing Tools: Entities Extraction

- **DBpedia Spotlight**
  - The Jena-Optronik ASTRO APS is an Autonomous Star Sensor with the most advanced radiation hard CMOS Active Pixel Sensor detector technology for long-term missions on Telecom, Science and Earth Observation satellites. Space for success.

- **OPEN CALAIS**
  - [Picture Alphasat © ESA](http://www.jena-optronik.com)
  - [Jena-Optronik GmbH](http://www.jena-optronik.com)
  - Phone: +49 3641 200-110
  - Fax: +49 3641 200-222
  - Email: sales@jena-optronik.de
  - Web: [www.jena-optronik.com](http://www.jena-optronik.com)

**Auto-Improvement of ontology**
**Introduction**

**Ontology**

**ConTrOn**

**Summary**

---

**Continuously Trained Ontology**

- **System Design Tools (e.g., CAD)**
- **Parts Database**
- **Ontologies**
- **Extracted Values from Data Sheets**
- **Class Names**
- **Augmented Entities**
- **External Semantic Knowledge Base (e.g., Wikidata)**
- **Lexical Database (e.g., WordNet)**
- **PDF Data Sheets**

---

**ConTrOn**

- **Highlighted PDF Data Sheets**
- **Highlighted Pattern Learner (KPL)**
- **Information Extractor (IE)**
- **Ontology Enricher (OE)**
- **Domain Knowledge Extractor (DKE)**

---

**Domain Expert**

- **Review Extracted Information**
- **Corrected Extracted Information**
- **Highlight Detected Values**

---

**Extraction Patterns**

- **Domain Representing Concepts**

---

**Key-Value Pattern Learner (KPL)**

---

**Information Extractor (IE)**

---

**Ontology Enricher (OE)**

---

**Domain Knowledge Extractor (DKE)**
Domain Knowledge Extractor

- Extract keywords
  - Bag-of-words
  - Tf-idf

- Word Disambiguation
  - https://wordnet.princeton.edu/
  - Part-of-Speech Tagging
  - Vector Space Model

Auto-Improvement of ontology
If ambiguous (multiple entities matched), compare to domain knowledge keywords.

At this step, only enriching the existing classes.
Information Extractor

- Search text based on ontology classes
  - Name, label, superclass, same as
- Extract values that come after keywords

<table>
<thead>
<tr>
<th>Extraction approach</th>
<th>Extracted property-value pairs</th>
<th>Recall</th>
<th>Precision</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideal</td>
<td>Ideal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ConTrOn</td>
<td>427</td>
<td>0.94</td>
<td>0.6</td>
<td>0.73</td>
</tr>
<tr>
<td>Text-based search</td>
<td>271</td>
<td>0.97</td>
<td>0.49</td>
<td>0.65</td>
</tr>
</tbody>
</table>

- Need human-in-the-loop: Next
Key-Value Pattern Learner

- Users provide feedback via a UI (human-in-the-loop)

- Key → Add to the ontology
  - Enrich by adding new classes

- Value → Improve the information extraction

Auto-Improvement of ontology
Pattern Example

• Learned patterns
  – `<number>"+" x "+<number>+" x "+<number> + <unit>`
  – `<number> <unit>+" x “+<number> <unit>+" x “+<number> <unit>`

• Verify the pattern
  – Apply to the information extractor
  – Choose the pattern that yield the minimum error
Summary

• Current satellite parts ontology is available at: https://zenodo.org/record/2616374

• ConTrOn: automatically improve ontology from data sheets
  ✓ Ontology Enrichment
  ✓ Domain Knowledge Extraction
  ✓ Ontology-based Information Extraction
  ❏ Key-Value Pattern Learner (human-in-the-loop)
Outlook

• Collecting feedbacks from users
  – Baseline for evaluation of information extraction

• Extension of ontology

• Knowledge graph from data sheets
Thank you for your attention!
Ontology: Concept

Ontologies do not only introduce a sharable and reusable knowledge representation, but can also add new knowledge about the domain.

https://www.ontotext.com/knowledgehub/fundamentals/
Semantic Knowledge

https://www.ontotext.com/knowledgehub/fundamentals/