The Multi-Intelligence Tools Suite (MITS)

Supporting Research and Development in Information and Knowledge Exploitation

Jean Roy
Group Leader – Sensemaking Group

Dr. Alain Auger

Intelligence and Information Section
Defence R&D Canada – Valcartier


21 June 2011
Outline

• Multi-Intelligence Tools Suite (MITS) definition
• Key characteristics of the MITS
• The MITS as a knowledge-based system
• Ingesting data into the MITS
• Overview of the MITS main components
• MITS exploitation example
• MITS implementation and evolution
Multi-Intelligence Tools Suite (MITS)

A federation of innovative, composable and interoperable intelligence related tools, which are integrated and interleaved into an overall, continuous process flow relevant to the intelligence community

Input

- Domain Specific Knowledge
  - Domain concepts
  - Inference Rules
  - Etc.
- Documents, Lists, Track Data, Etc.

Output

- Facts
- Targetted information

MITS
- Document Management
- Text Document Processing
- Semantic Analysis
- Automated Annotation
- Fact Extraction
- Trend Analysis
- Manual Fact Insertion
- Automated Reasoning
- Subject List Processing
- User Notification
- Etc.
Key Characteristics of the MITS
Multi-Intelligence Tools Suite (MITS) Key Characteristics

- **Multi-Int Analysis**
  - I&I Section primary focus
  - HUMINT, SIGINT, IMINT, OSINT, Etc.
  - Intelligence products (no “raw” data)
- **Tool Suite: One-Stop Shop for the Tools**
  - Easy access to all available tools
- **Tool Synergy**
  - Mutually advantageous conjunction and compatibility
- **Ease of Exploitation**
  - Facilitates the transfer of data, information and knowledge
  - Seamless pipelining of the individual tools
- **Uniformity of Exploitation**
  - Organization of knowledge
  - Standardized HCI / GUI

- **An R&D Integration Environment**
  - Unconstrained environment
    - No imposed technologies
  - Reusability and incremental development
    - Spiral development approach
    - Not having to start from scratch every time
  - Integrated validation
    - Already validated components provide realistic stimulation
    - Overall components interactions
- **Knowledge-Based Systems Technologies**
  - Generic technology (“easy” to adapt to a new domain)
  - Domain dependent analysis (MDA, IED, terrorism, etc.)
  - Cross-domain analysis
The MITS as a Knowledge-Based System
Knowledge-Based System Technologies (1/2)

« Processing is separated from the problem-solving knowledge »

- Represent knowledge in a more natural fashion
- Focus is on capturing and organizing problem-solving knowledge
- Allow changes to be made to the knowledge base without side effects on program code
- Allows the same control and interface software to be used in a variety of systems, in different domains
- Allows to experiment with alternative control software for the same knowledge base
Knowledge-Based System Technologies (2/2)

- Generic
- Developed “only once”
- Developed by “others”:
  - Comercially available
  - Open source

- Specific to a domain
  « Knowledge Cartridge »
- Requires knowledge acquisition from domain experts
  « Very difficult
- Requires knowledge representation
  « Knowledge Engineering »
Multi-Intelligence Tools Suite (MITS)
A Knowledge-Based System

• **Knowledge Representation Building Blocks**
  – Ontologies
  – Facts
    ‣ Atom definitions
  – Built-in definitions
  – Inference rules
  – Text-based templates (for fact extraction)

• **Domain Knowledge (A Priori / Reference / « Static »)**
  – Domain environment knowledge
    ‣ Ontologies and taxonomies
  – Domain expert knowledge (know-how)
    ‣ Inference rules
    ‣ Text-based templates
  – Utilities
    ‣ Local grammars (pattern matching rules)
    ‣ Sources characterization
    ‣ Fact generators

• **Situation Knowledge Base (Dynamic Situation Model)**
  – Situational facts
  – Situational ontologies

• **Knowledge Engineering Module**
Domain Environment Knowledge (MITS)
(Ontologies / Taxonomies)

Class: Represents a concept
Individual: Represents an object

Knowledge Domain
Situational Facts and Atom Definitions

A fact is a pragmatic « truth », a statement that can, at least in theory, be checked and confirmed.

« Ship X is in proximity of infrastructure Y »

Fact in
Natural Language

In Proximity (Ship X, Infrastructure Y)

Fact in
Formal Language

In Proximity (Ship Name, Infrastructure Name)

Atom Definition
(Fact Structure Template)
Situational Facts: Atom Definitions (Fact Formal Structure Template)

- Represents a template (the formal structure) on which facts in the system will be based.
- It is defined by a **name** and a **list of arguments** with a precise **type** and **order**.

```
A. Atom definition name

Member of (Bin Laden, Al-Qaeda)
```

Atom Definition Could be Used to Create a Fact

Member of (Bin Laden, Al-Qaeda)
Built-In Definitions

• Can be used in inference rule premises
• Defines functions that can be invoked to perform custom validations during the inference process
  – Evaluate their arguments using functions pre-embedded in the system
  – Returns « True » or « False »
• The MITS is pre-initialized with a predefined set of built-in definitions
  – They cannot be edited, deleted or neither a new built-in definition can be added
• The current set:
  – after
  – before
  – collect
  – inArea
  – includeDate
  – includeDateRange
  – overlap

Validates if a geometry (e.g., a point) is contained within a reference area
Validates if the ranges of two dates overlap
Inference Rules

- Used by the rule-based inference functionality to infer new facts
- Defines which pattern of facts will generate new facts
  - **Inference rule premises** (the « IF » part)
    - Define the facts required to be validated as « True » in order to trigger the rule
    - Composed of atom definitions, built-in definitions and operators
  - **Inference rule conclusions** (the « THEN » part)
    - Define the facts that will be created when the rule premises are satisfied (i.e., when facts matching the premises are found in the system)
    - Composed of atom definitions and operators
Text-Based Templates

• To find precise series of words in unstructured text documents (Word, PDF, etc.) and to extract specific facts from them

• Text-based « processing rule » defining
  
  − **Constraint**: A text-based pattern to search for in text documents
   
   ▶ Find precise series of words in a document
   
   ▶ The user can define:
     
     ➢ The order of words he/she is looking for
     
     ➢ If the template has to match with an instance, instance type, or instance type pattern (of a knowledge domain ontology)
     
     ➢ A word’s lexical category
  
  − **Conclusion**: The facts that will be created when this text-based pattern is found

```plaintext
Constraints

[ ( Person name ), ( A was , A seen , A in ), ( Afghanistan ) ]

Conclusions

Has been seen at ( group1 , Afghanistan , A = 2009/11/05 ) + asdt ( ? group3 )
```
Knowledge Representation Building Blocks and Situational Facts Generation

- Domain Knowledge Ontologies
  - Used By
  - Used for

- Text-Based Template
  - Used By

- Semantic Analysis of Text Documents: Automated Fact Extraction
  - Generates
  - Used By

- Structured Data/Information Import (GPW Type)
  - Generates
  - Used By

- List Import Module
  - Generates

- Manual Fact Insertion
  - Generates

- Rule-Based Inference Module
  - Generates/Uses
  - Used By

- Atom Definitions
  - Are Similar To
  - Used By

- Inference Rules
  - Used By

- Built-In Definitions
  - Used For

- Situational Facts Knowledge Base
  - Generates
Ingesting Data into the MITS
Ingesting Data into the MITS

- **Instance Types**
  - **Instances**
  - **Observed Facts**
  - **Inferred Facts**

- **Application Domain Knowledge Base**
  - **Ontologies**
    - **Instance Types**
    - **Instances**
  - **Situational Facts Database**
    - **Facts**
      - **Observed Facts**
      - **Inferred Facts**

**Sources**

- **Structured Sources**
  - $S_1$
  - $S_2$
  - $S_n$

- **Unstructured Text Documents**
  - $S_1$
  - $S_2$
  - $S_n$

**Multi-Intelligence Tools Suite**

- **Source Adapter/Converter**
  - Automated Extraction
Overview of the MITS Main Components
Multi-Intelligence Tools Suite (MITS)

Features in a Nutshell

• Natural language processing capabilities to support automated semantic analysis of unstructured documents

• Automated place name disambiguation and geo-referencing of any piece of information

• Semantic and geospatial search for information in sources

• Automated entity extraction

• Automated collation of all entities of interest pertaining to the knowledge domain(s) of interest:
  − Person names, date and time elements, locations, organizations, components, effects, triggering mechanisms, types, etc.

• Automated/manual fact extraction capabilities from observations contained in sources (published intelligence products in general)

• Automated reasoning capabilities over facts

• List-based processing

• Support to trend and pattern analysis

• Automated and personalized alerting/notification capabilities
Multi-Intelligence Tools Suite (MITS)

Modules

- Structured Data/Information Import
  - Track Modeling
  - Data/Information Preview
  - Data/Information to Fact Conversion
- Unstructured Text Documents Processing
  - Document Repository / Management
  - Document Viewer
  - Semantic Analysis
    - Automated Annotation
    - Automated Fact Extraction
    - Geo-Referencing
  - Statistical Analysis of Annotations
    - Trend Analysis
- Manual Fact Insertion
- Automated Reasoning / Inference
  - Rule-Based, Case-based, Description Logic, Kinematics & Geospatial Analysis
- Subject Lists Processing
  - List-Based Situation Monitoring
  - Situation-Based List Filling
- Administration
  - Server Load, Inference, Ontologies
- Situational Fact Management
  - Situational Fact Knowledge Base
    - Fact Viewer (Tabular Format)
    - Fact Filtering
    - Fact Export (KML Output)
- Personalized User Notification
  - Notification Subscription
    - Fact Notification
    - Document Notification
  - Notification Management
- Knowledge Engineering
  - Atom Definitions
  - Built-In Functions
  - Text-Based Templates
  - Inference Rules
  - Domain Knowledge Base
    - Ontology Browser
- GIS
  - Area Manager
- Online Help
- User Login
MITS
Exploitation Example
MITS Exploitation Example (1/6)

• Initial Trigger
  – A rather vague input from a source
    – There’s going to be a drug related event involving a ship
    – Expected to happen in some (rather large) area at sea
    – In a given time window
    – Involving a person listed on a list of suspects

• Analysis Problem:
  – Find a vessel that:
    1. Has a historical record of drug smuggling
    2. Is associated with a suspect person designated by Agency X on a list
    3. Is on a list of suspect vessels provided by external Agency Y
    4. Is own by a person who currently has serious financial problems
    5. Is within the area where the drug smuggling event is expected to happen in the given time window
MITS Exploitation Example (2/6)

List of Suspect Persons Designated by External Agency X

Unstructured (Text) Documents (Relevant Collection)

Document Repository / Management

Semantic Analysis of Text Documents: Automated Fact Extraction

List Import Module

List of Vessels Associated with Suspect Persons

List of Vessels with Drug Smuggling History

List of Suspect Vessels

Manual Fact Insertion

Manual Fact Insertion

Owner Z (of Many Vessels)

« Has Financial Problem » Facts

List of Persons with Financial Problems

List of Persons with Financial Problems

Structured Data/Information Import (GPW Type)

«Vessel Facts» From Track Data

Area Manager

Area Manager

Smuggling Event Expected Area

Track Data In GPW Format

Inference Rules

Personalized User Notification

Automated Reasoning (Rule-Based Inference)

« On List » Facts

Notification Subscription

Alert Notification

« Alert » Fact

MITS Exploitation Example (2/6)
### MITS Exploitation Example (3/6)

<table>
<thead>
<tr>
<th>Derived With MITS</th>
<th>Derived With MITS</th>
<th>Externally Provided</th>
<th>Derived With MITS</th>
<th>Derived With MITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Vessels Associated with Suspect Persons</td>
<td>List of Vessels with Drug Smuggling History</td>
<td>List of Suspect Vessels</td>
<td>Vessels in Smuggling Event Expected Area</td>
<td>Vessels Having Owner in Financial Trouble</td>
</tr>
<tr>
<td>KING GRATTON</td>
<td>BRUTUS</td>
<td>ZIPPO POTOMAK</td>
<td>STRATOCASTER</td>
<td>TATIANA</td>
</tr>
</tbody>
</table>
List of Vessels Associated with Suspect Persons

List of Vessels with Drug Smuggling History

Vessels in Smuggling Event Expected Area

Vessels Having Owner in Financial Trouble
Rule #2

Has Financial Problems (Person Name) → Owner Has Financial Problems (Vessel Name)

Owner (Vessel Name, Person Name)

Rule #3

In Area (Vessel Name, 2D Coordinates, Reference Area) → Alert (Vessel Name)

Position (Vessel Name, 2D Coordinates)

On List (Vessel Name, List of Vessels Associated with Suspect Persons)

On List (Vessel Name, List of Suspect Vessels) → On List (Vessel Name, List of Common Vessels)

On List (Vessel Name, List of Vessels with Drug Smuggling History)

Rule #1
MITS Exploitation Example (6/6)
Notification – Justification
MITS
Implementation and Evolution
MITS Implementation and Evolution (1/2)

Intelligence Science and Technology Integration Platform (ISTIP)
Visionary Overarching Interaction Interface Layer for the Analyst (VOiiLA)
MITS Implementation and Evolution (2/2)

Intelligence Science and Technology Integration Platform (ISTIP)
Visionary Overarching Interaction Interface Layer for the Analyst (VOiILA)
Questions ???