



Automating Requirements Analysis

Using
Natural Language Processing
And Semantic Knowledge

Why Use NLP On FPS?

An FPS is foundational – there is high value in getting it right, and having everyone read it the same way

The language is formal and technical

Altering a few words is much easier than rework of built equipment

Mere analysis, inspection, or review alone will find some of these issues but *generally is far weaker than usually is realized* - Wikipedia

4 Pieces

People can handle no more than four pieces of information in play at once – more than that and they chunk them

Overload

4

4

Pieces

We know what happens when a person is overloaded – they make mistakes

Flying a helicopter is near the limit of what a person can do – a mistake can cause an immediate crash

Overloading a person with a new specification can cause a crash – it is just delayed a long time, so harder to see the cause



We can build a model of it


We already have a model of it
– in natural language
Why not use it?

what is necessary

4
4
Pieces

Documents as Machinery

The goal is to have documents treated as just another piece of machinery. Then we can have tools to tell us about:

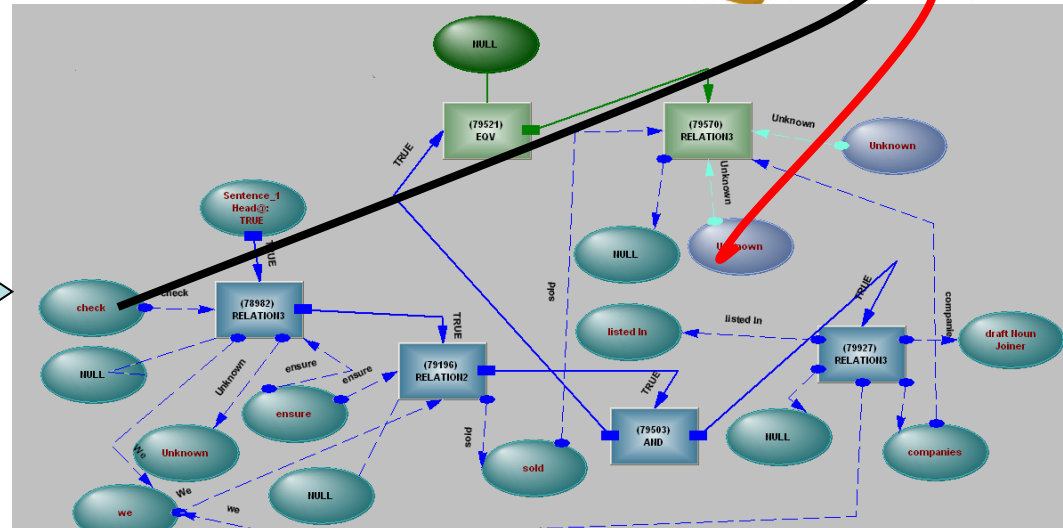
- Things that are broken, not connected, or assembled the wrong way
 - A component not working properly
 - Something missing – not logically complete
 - Confusion from too many pieces in play
- 

This is much easier to do on a realised working structure because the structure does most of the work



3.2.4.3 Chapter 1, Space System Description. This chapter shall describe the space system and supporting facilities in sufficient detail to afford the reader a single source document of general system information. A brief narrative shall describe the purpose, main features and leading particulars for the space system as described below. Illustrations that clarify a particular system or reduce the verbiage necessary for explanation shall be included.

3.2.4.3.1 Description of the satellite. The description shall be of sufficient detail to provide an understanding of the purpose and function of the subsystems, their relation to overall system operation, and such additional information as to enable the crew member to understand subsystem functions peculiar to the



The Bad News

A specification is going to be a bit more difficult than project management or stress analysis, where simple approaches work

- It defines its own terms, it refers to itself and to other documents, it turns parts of itself on and off
- It can bring in existence, relations on relations, groups – anything natural language allows
- It needs to be disambiguated in context

Solving only one facet of a complex problem can be worse than useless

The Good News

A Function and Performance Specification is intended to:

- Describe what is required
- Be unambiguous
- Be consistent and coherent
- Be complete
- Use “dumbed down” language – only the “shalls” count

Sometimes it is necessary to fall back on the freedom of natural language. It may only take a few words, but those few flashes of complexity destroy the usefulness of any narrowly focused approach.

Building Document Structures



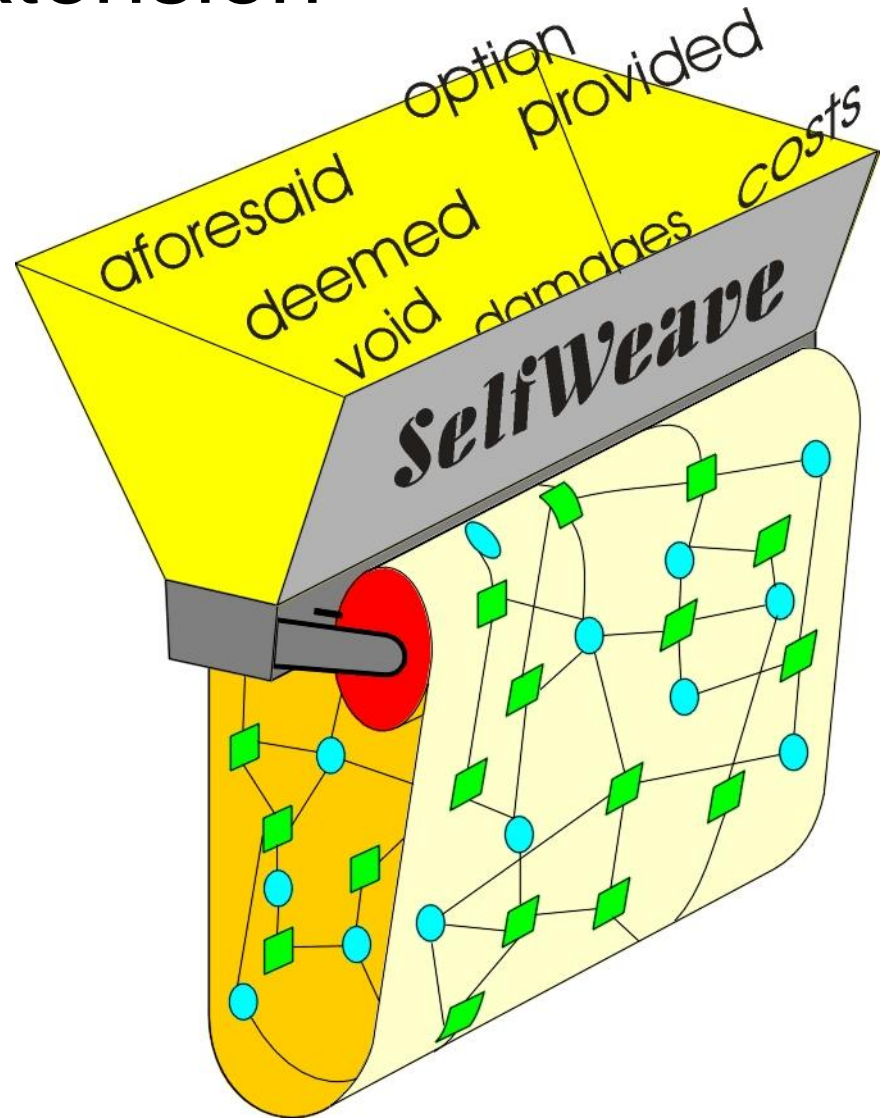
A large document may use a million small undirected elements in its structure.

We can try to build these large complex structures by hand, but then we run into our four pieces limit, or we can build a machine which builds structures - just like a Jacquard loom, and for the same reason



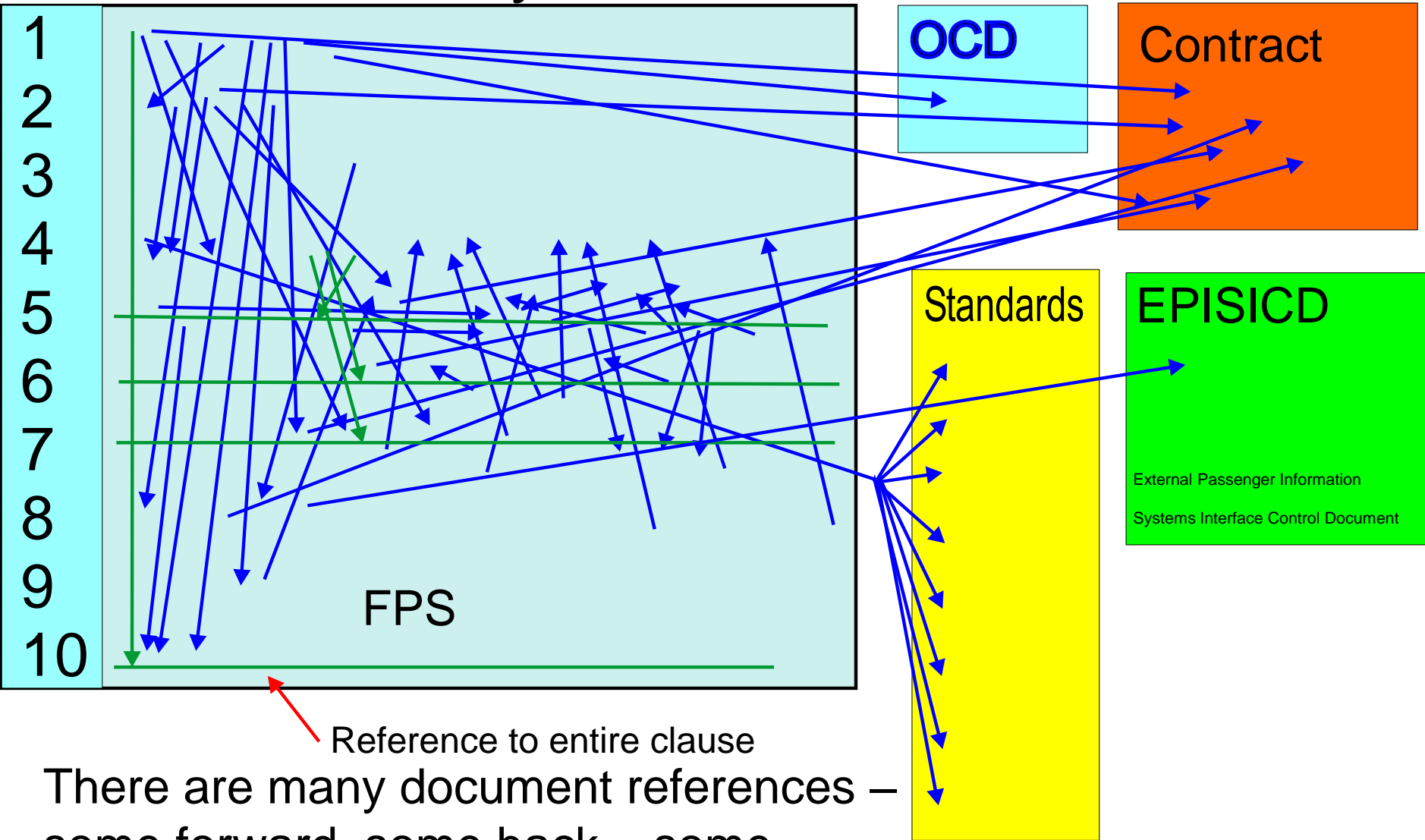
Self Extension

We need a machine which uses an initial structure to read text, extends its structure as it does so, then uses that extended structure to read more of the text – the way a person does



“Available Hours” means the period defined in clause 4.3.1.1.
TGRS Line Replaceable Unit (LRU) replacement shall be....

SubSystem Connections



Reference to entire clause

There are many document references – some forward, some back – some precise, some not - “Clause 10”.

FPS

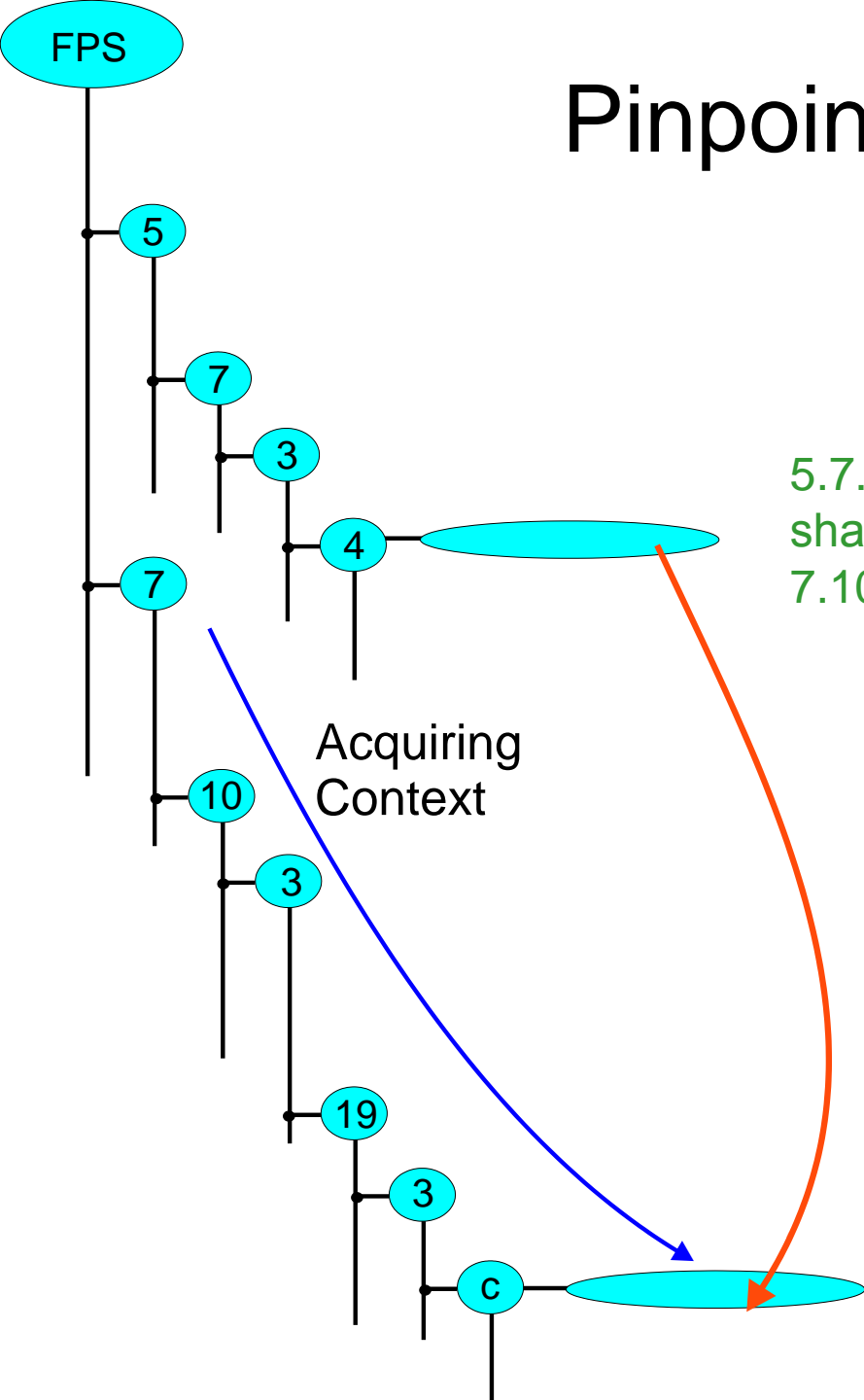
Pinpoint Precision

5.7.3.4. The On-Board Audio Announcer shall provide ...as described in clause 7.10.3.19.3.c

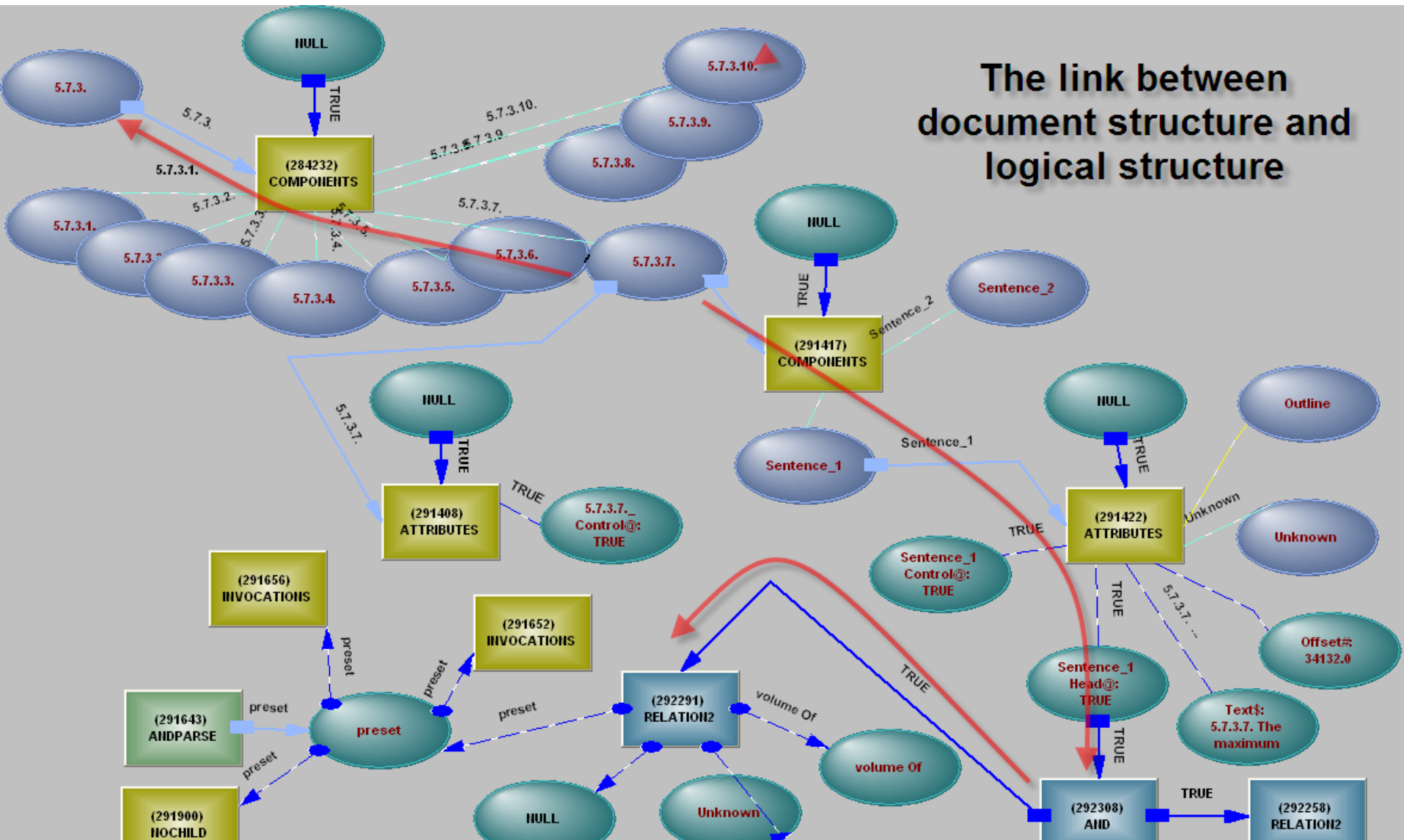
Acquiring Context

When reading Clause 5, Clause 7 hasn't been seen yet – the reference is logged for later validation and connection

The reading system spends a lot of time scheduling itself – jobs at the end of a sentence, at the end of a clause, at the end of reading



Document and Text



The system needs to be able to find its way from the relation structure in one part of the FPS to relation structure in another part, through a document reference

Tow Motor Example



Even at this low level of complexity (not much), automated FPS analysis can show errors, inconsistencies, ambiguities, duplications, weak statements

The resulting semantic structure can be shared between client and contractor, and used to support the entire process

Policies are just more text machinery, so they can be mixed in

Weak Statements

5.105 The Tow Motor shall be designed to the human engineering requirements of SAE 1247C.

....

5.110 The Tow Motor shall have human engineering design requirements that conform to SAE ARP 1247C.

There is duplication five lines later, but –

These look the same, but one is more restrictive than the other – the second one is so weak as to be meaningless

Errors

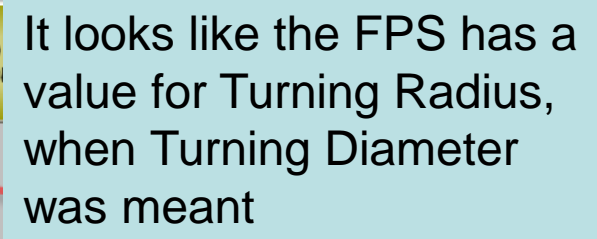
5.85 The construct of a Tow Motor shall ensure the ingress of air-borne dust and sand into the cabin.

The wrong end of the stick. We have a definition of a comfortable environment - 15 - 25° C, 20-70 dBA, a maximum particulate level – increasing the particulate content is inconsistent with that, but having a cabin heater is not

The turning radius of the Tow Motor shall not exceed 15 m.

Seems large – turning circle diameter? – truck typically 6-10m radius

"Turning Radius"



Ambiguity

There can be ambiguity within sentences, at conjunctions, at words with multiple meanings, at prepositions, at packed noun phrases.

And then there is ambiguity over objects many sentences apart:

Emergency brake interlock system shall be provided to prevent the movement of the vehicle with the brake applied.

Which brake? – there is a parking brake and a service brake. The system will see the ambiguity and ask the question – the writer can add a word to eliminate the ambiguity.

Often these problems only need the pieces of text to be brought together on the same page to be obvious - “just show me the stuff about braking”

The writer is usually the last person to see the ambiguity, because they know what is meant.

Reporting Ambiguity

The screenshot shows the FPS Parser application window. The title bar reads "FPS Parser - [E:\Orion\Example\TowMotorSave.HTM]". The menu bar includes File, Edit, Options, Tools, Windows, and Help. The toolbar contains various icons for file operations and editing. The main window has tabs for Input Text, Tokeniser, Tag List, Scratch Pad, Query, Current Paragraph, and Reporting. The Reporting tab is active, displaying a report titled "Report" with sub-sections "Report Ambiguity" and "Ambiguous Child". A list of report types is visible on the left, including ReportAmbiguity, MandatoryAttribute, ForbiddenRelation, InvalidReference, InconsistencyWithinDocument, InconsistentWithPriorKnowledge, WeakStatement, RelationDuplication, PoorPhrasing, and SemanticIndecision. A "Prepare Report" button is at the bottom left. A text box on the right explains that ambiguity can arise from a word having multiple meanings, a conjunction, a collocation, or a packed noun phrase. The main text area shows an example of ambiguity from a document, highlighting the word "brake" in different contexts.

There can be ambiguity at a word having several meanings, across a conjunction or at a collocation or a packed noun phrase, or there can be global ambiguity – where several statements need to be seen at once, as in this example

Report

Report Ambiguity

Ambiguous Child

Ambiguity from brake To parking brake And service brakes

In: 5.102 An emergency brake interlock system shall be provided to prevent the movement of the vehicle with the brake applied

To 5.15 The Tow Motor shall have a parking brake that shall engage on at least two (2) wheels and shall hold a fully ballasted Tow Motor in neutral shift , at standstill , on a minimum incline of 5 degrees (as detailed in AHM 913) or maximum slope for ramp areas as published in ICAO annex 14 Aerodromes Volume 1 - Aerodrome Design and Operations , whichever is the greater

And 5.18 The Tow Motor shall be equipped with a warning indication for a malfunction of the service brakes

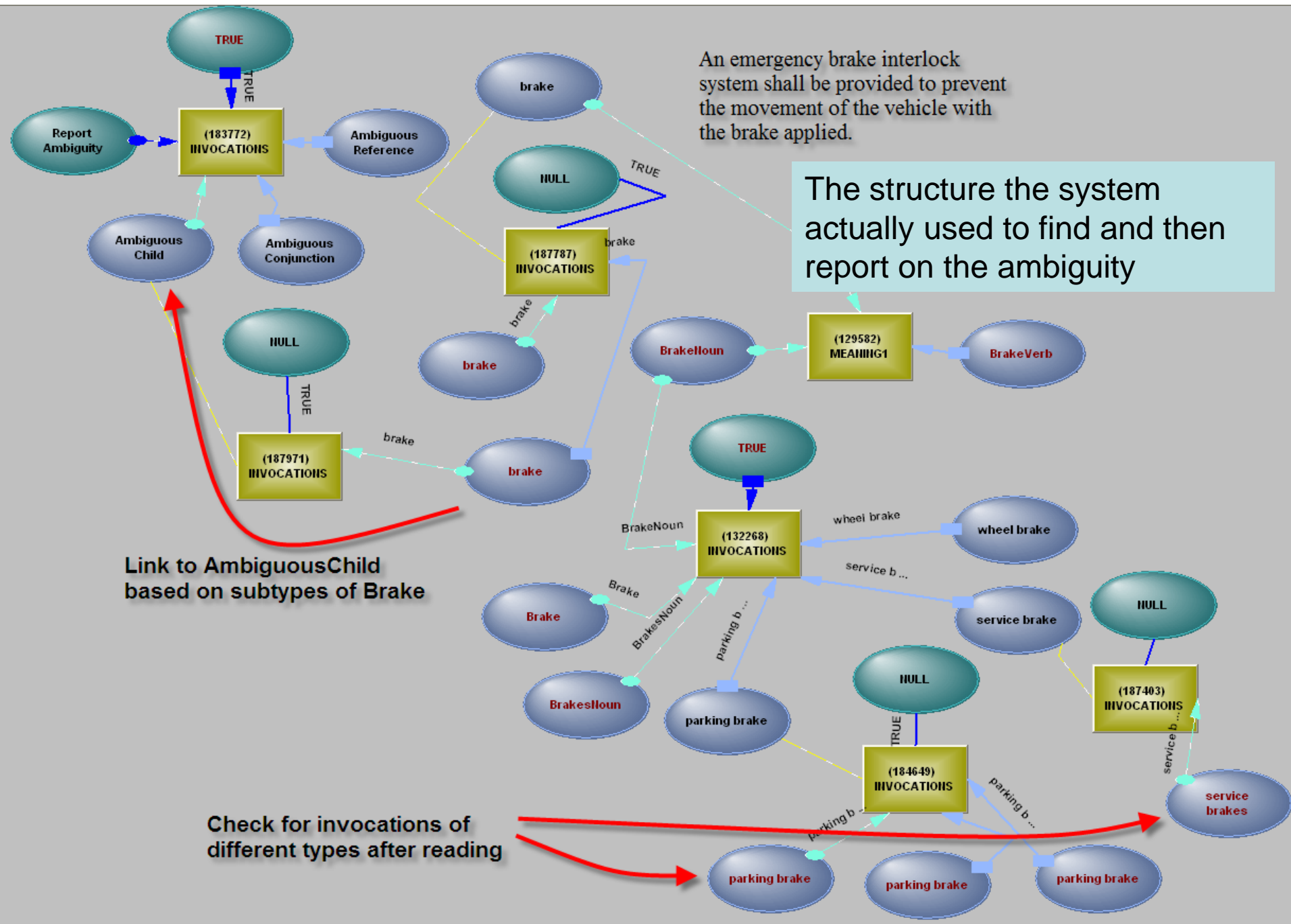
Read E:\Orion\Example\TowMotorSave.HTM FILE STREAM E:\Orion\Example\TowMotorSave.HTM

An emergency brake interlock system shall be provided to prevent the movement of the vehicle with the brake applied.

The structure the system actually used to find and then report on the ambiguity

Link to AmbiguousChild based on subtypes of Brake

Check for invocations of different types after reading



Semantic Search

The system can read, so it can do Semantic Search. You can enter a search string without worrying about knowing the exact way something is described in the specification – the meaning is searched for, not the words

Semantic Search will proceed through all synonyms, false antonyms, mapped relations, packed noun phrases

You can ask for powerful searches using relations

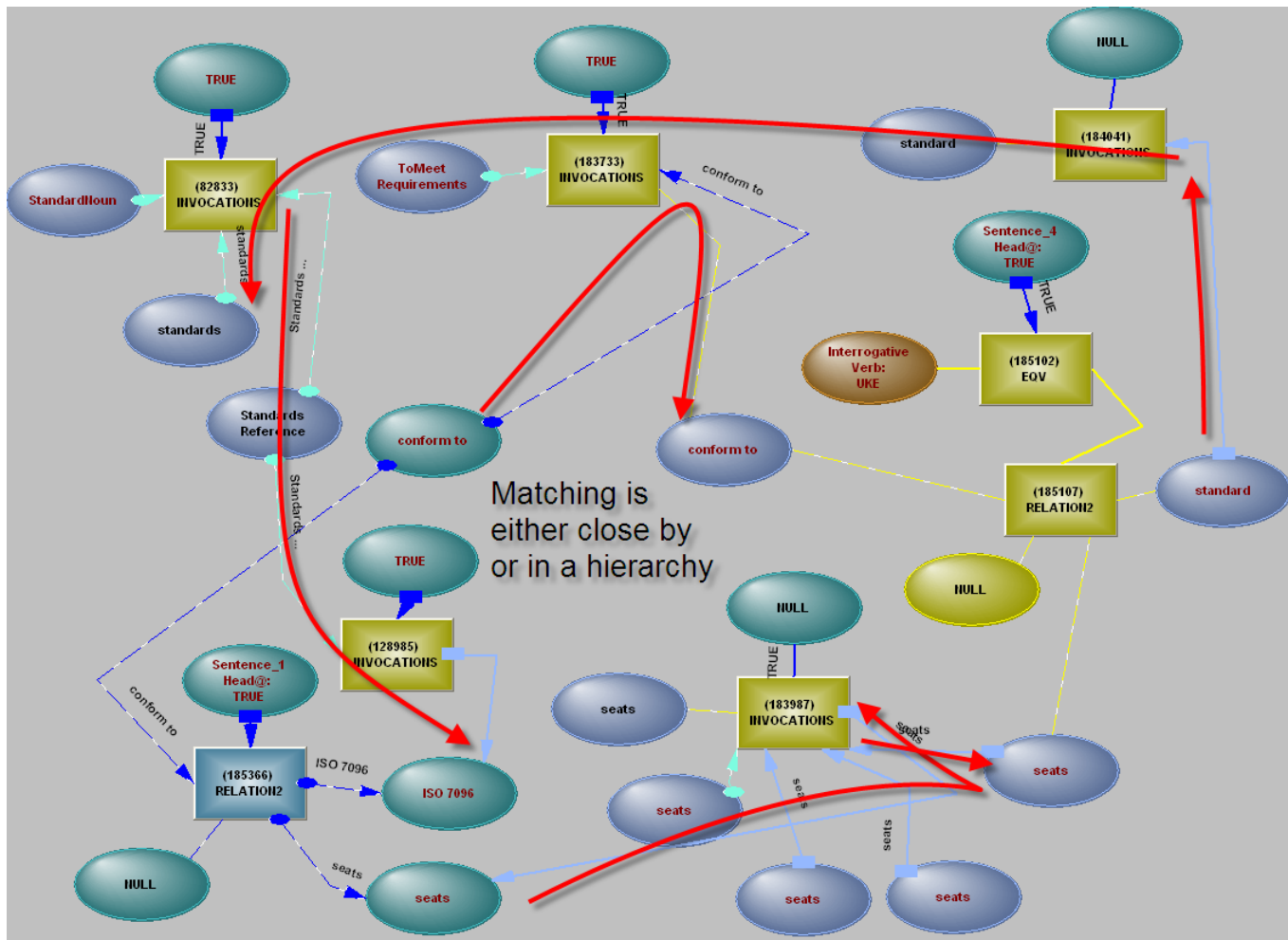
Do the Tow Motor brakes comply with a standard?

What instrumentation is not located in the cabin?

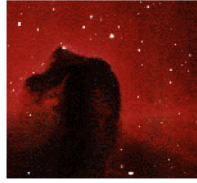
You are querying a more complex structure than you can hold in your head – you have broken your four pieces limit

Searching

A query structure is created using free text, and it is then matched to the structure built from the specification. These aren't two different structures – the query structure is anchored in the structure built from the text to be searched, so common points are found during building of the query.



- Treat documents as pieces of machinery
- Make all the connections before beginning requirements analysis, so all possible influences can be seen by the system
- Accept that people have a limit on pieces of information in play – don't fight it, work out how to get around it
- If it is too big to build in your head, build it directly from the text, rather than building it physically, and make sure you include everything, not just the easy bits



Requirements Engineering Using Active Structure

Australia
Interactive Engineering Pty Ltd
www.semanticstructure.com