Evaluating OWL 2 Reasoners in the context of Clinical Decision Support in Lung Cancer Treatment Selection

M. Berkan Sesen Ernesto Jiménez-Ruiz
René Bañares-Alcántara Sir Michael Brady

Department of Engineering Science
Department of Computer Science
Department of Oncology
University of Oxford, UK

2nd OWL Reasoning Evaluation Workshop
22 July 2013

Outline

Background

Lung Cancer Treatment

- Lung cancer is responsible of the 21% of cancer-related deaths.
- There are (substantial and unjustified) variations in treatment decisions between cancer centres.
- Clinical guidelines (CGs) reduce variability in clinical practice.
- Originally CGs are unstructured and free-text documents, and often not readily accessible at the point of decision making.

Background

Lung Cancer Treatment

- Lung cancer is responsible of the 21% of cancer-related deaths.
- There are (substantial and unjustified) variations in treatment decisions between cancer centres.
- Clinical guidelines (CGs) reduce variability in clinical practice.
- Originally CGs are unstructured and free-text documents, and often not readily accessible at the point of decision making.

Background

Lung Cancer Treatment

- Lung cancer is responsible of the 21% of cancer-related deaths.
- There are (substantial and unjustified) variations in treatment decisions between cancer centres.
- Clinical guidelines (CGs) reduce variability in clinical practice.
- Originally CGs are unstructured and free-text documents, and often not readily accessible at the point of decision making.

Background

Lung Cancer Treatment

- Lung cancer is responsible of the 21% of cancer-related deaths.
- There are (substantial and unjustified) variations in treatment decisions between cancer centres.
- Clinical guidelines (CGs) reduce variability in clinical practice.
- Originally CGs are unstructured and free-text documents, and often not readily accessible at the point of decision making.

Background

Clinical decision support (CDS) systems can...

- facilitate the access to clinical guidelines.
- computerise CGs using structured logical languages.
- match guidelines rules against a patient record to infer the appropiate treatment.

Examples

- PROforma. Fox et al. (1997)
- EON. Musen et al. (1996)
- GLIF3. Want et al. (2004)
- SAGE. Tu et al. (2007)
- **LUNG CANCER ASSISTANT**. Berkan Sesen et al. (2012)

Background

Clinical decision support (CDS) systems can...

- facilitate the access to clinical guidelines.
- computerise CGs using structured logical languages.
- match guidelines rules against a patient record to infer the appropiate treatment.

Examples

- PROforma. Fox et al. (1997)
- EON. Musen et al. (1996)
- GLIF3. Want et al. (2004)
- SAGE. Tu et al. (2007)
- **LUNG CANCER ASSISTANT**. Berkan Sesen et al. (2012)

Background

Lung Cancer Assistant (LCA)

- An **ontology-based** system which provides guideline rule-based decision support for lung cancer treatment.
- LCA exploits the English Lung Cancer Dataset (**LUCADA**)

LUCADA ontology

- LUCADA has been built using the OWL 2 language.
- Represents the semantic layer of the LCA:
  - Captures the **domain** in the LUCADA dataset.
  - Encodes the **clinical guidelines**.
  - Represents **patient data**.

Background

Lung Cancer Assistant (LCA)

- An **ontology-based** system which provides guideline rule-based decision support for lung cancer treatment.
- LCA exploits the English Lung Cancer Dataset (**LUCADA**)

LUCADA ontology

- LUCADA has been built using the OWL 2 language.
- Represents the semantic layer of the LCA:
  - Captures the **domain** in the LUCADA dataset.
  - Encodes the **clinical guidelines**.
  - Represents **patient data**.

Outline

LUCADA Ontology

Example of guideline rule

- Eligibility criteria are encoded as equivalence axioms.
- *"Consider radiotherapy for Stage I, II, III patients with good performance status"*

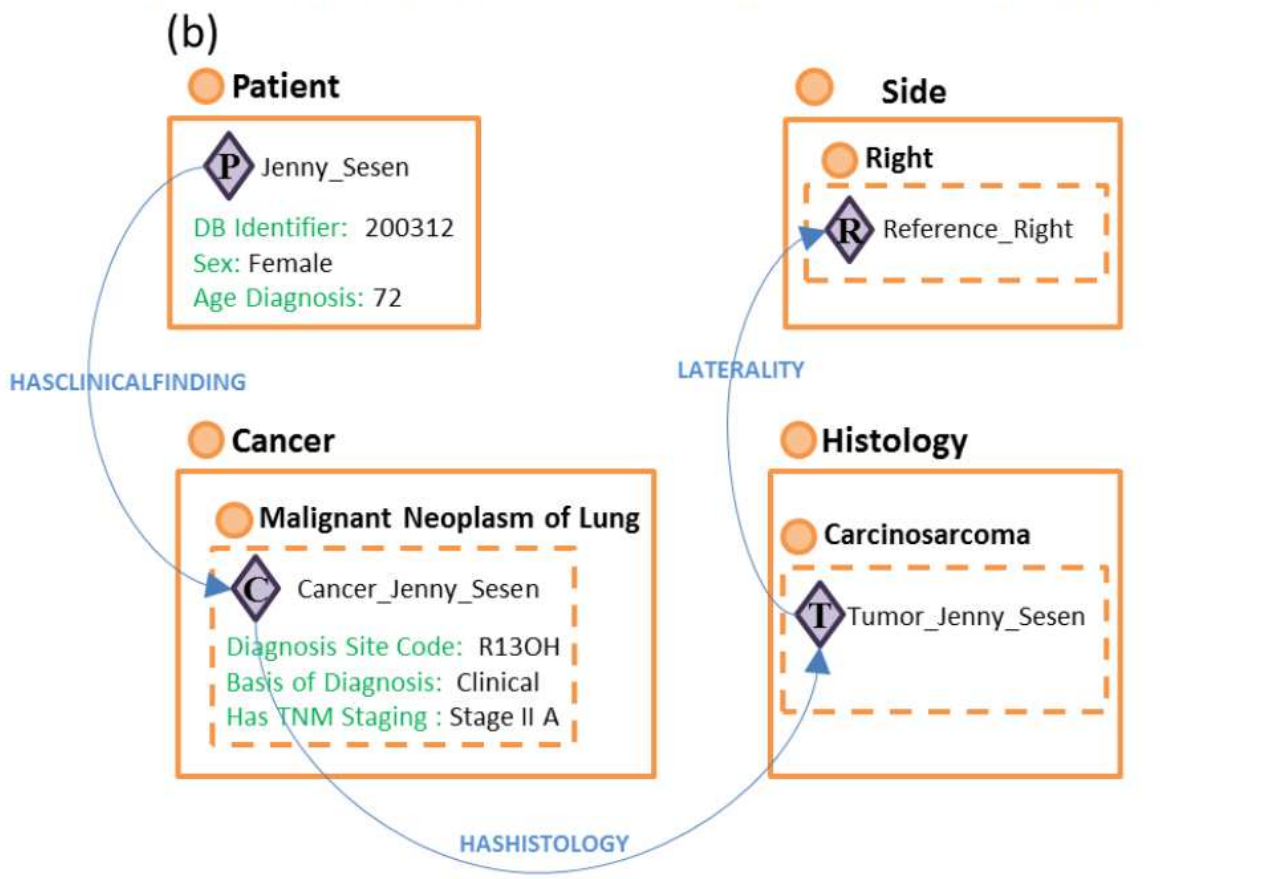
$$\begin{aligned}\mathsf{RT\_GR} \equiv\ & \mathsf{GoodPerformancePatient} \sqcap \exists \mathsf{hasClinicalFinding}. \\ & (\mathsf{NeoplasticDisease} \sqcap \\ & \exists \mathsf{hasPreHistology}.\mathsf{NonsmallCellCarcinoma} \sqcap \\ & \exists \mathsf{hasPreTNMStaging}.\mathit{string} \sqcap \\ & \forall \mathsf{hasPreTNMStaging}.\{\mathit{I}, \mathit{II}, \mathit{III}\})\end{aligned}$$

LUCADA Ontology

Example of patient

- Each patient is encoded with $\sim$ 25 individual axioms.

DB Identifier	Name	Sex	Age	Primary Diagnosis	TNM Staging	Histology	Tumour Laterality
200312	Jenny_Sesen	F	72	Malignnant Neoplasm of Lung	II A	Carcinosarcoma	Right



LUCADA Ontology

Integration with SNOMED CT

- SNOMED is the reference ontology in the National Health Service (NHS).
- To facilitate interoperability we have integrated LUCADA with SNOMED.
- We have used LogMap matching system to
  - identify the classes in SNOMED related to LUCADA.
  - extract a lung cancer-specific module of SNOMED CT.

LUCADA Ontology

Summary of LUCADA and LUCADA-SNOMED metrics

Metric \ Ontology	LUCADA-SNOMED	LUCADA
DL Expressivity	$\mathcal{ALCHIF}(\mathcal{D})$	$\mathcal{ALCHI}(\mathcal{D})$
# Classes	1553	376
# Object properties	63	37
# Data Properties	63	63
# Equiv. class axioms	1050	40
# Subclass of axioms	999	386
# Prop. domain axioms	97	97
# Prop. range axioms	30	30

Outline

Evaluation

Evaluation settings

- Windows 7 64-bit desktop computer,
- 15 GiB of RAM, and
- Intel Xeon 2.27 GHz CPU.
- Results have been calculated as average of at least 10 repetitions of the experiment.

Evaluation

Evaluated Reasoners

- HermiT 1.3.7, Pellet 2.3.0 and FaCT++ 1.6.2

Experiments

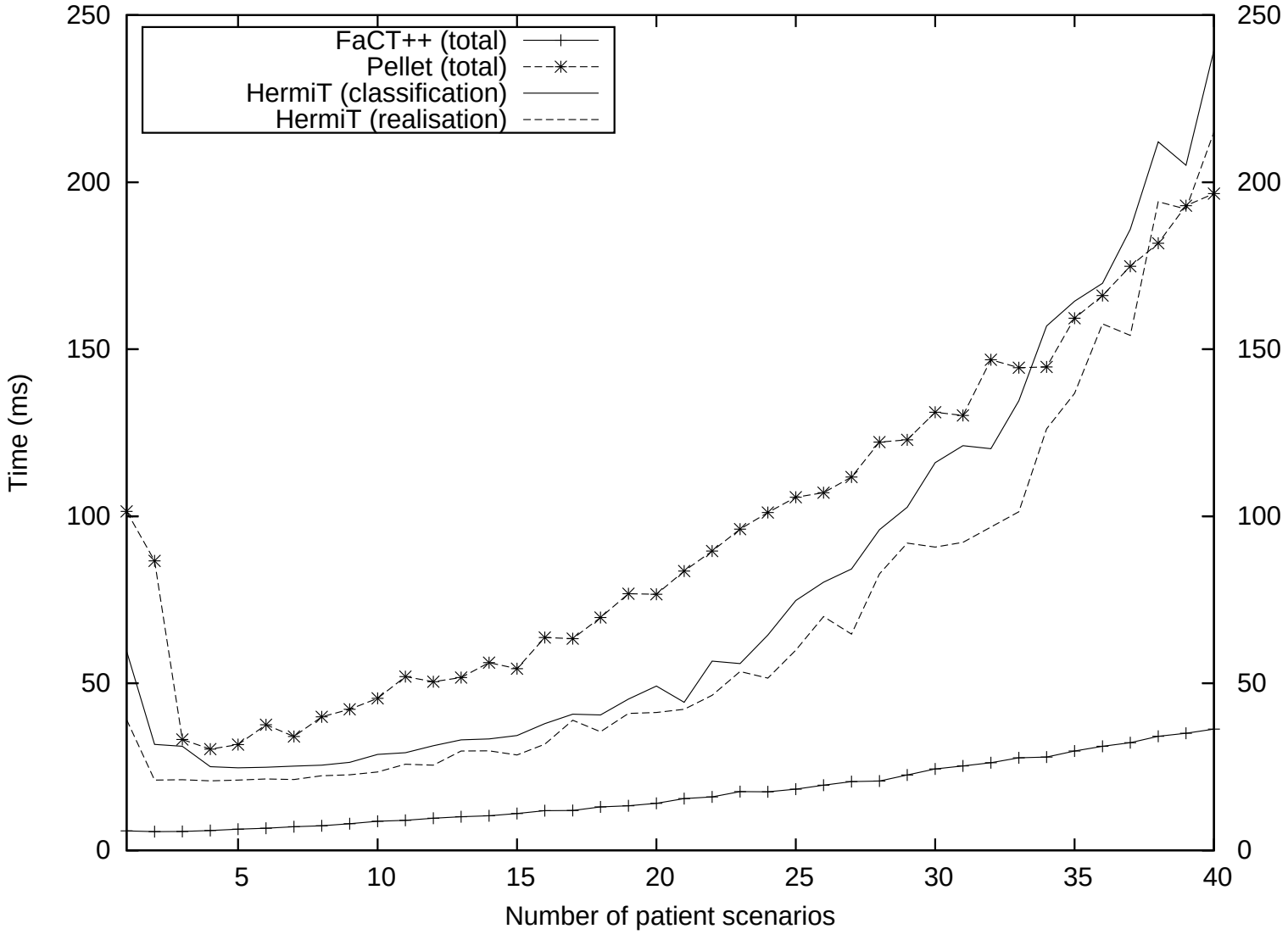
- Increasing the TBox with guideline rules or patient scenarios.
- Increasing the ABox with patient records.

Evaluation

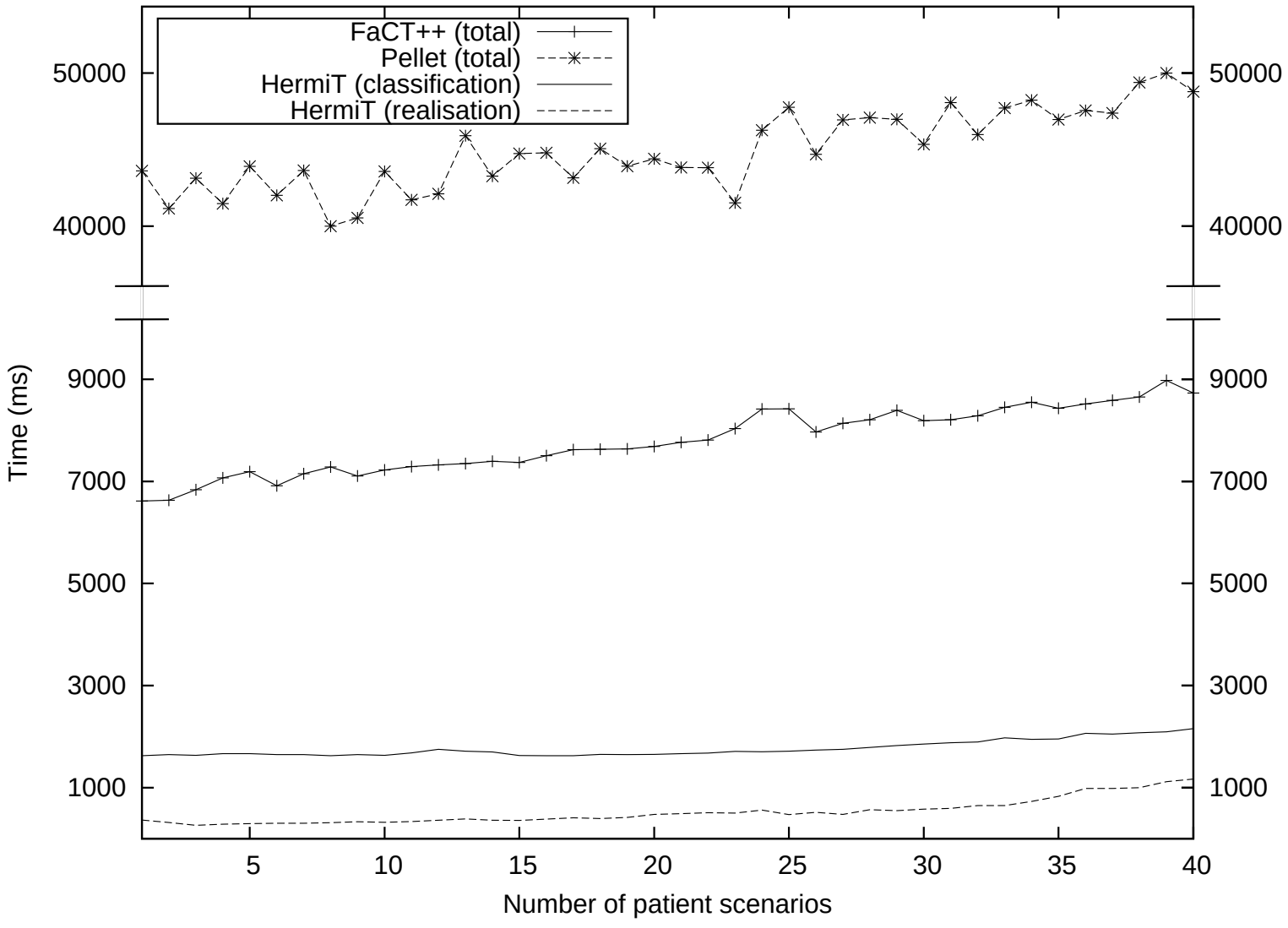
Experiment 1: Increasing the TBox with guideline rules

- 1 to 40 patient scenarios or guideline rules.
- With LUCADA and LUCADA-SNOMED with 1 patient.
- Recorded times for classification and realisation.

Experiment 1 (increasing TBox) with LUCADA



Experiment 1 (increasing TBox) with LUCADA-SNOMED

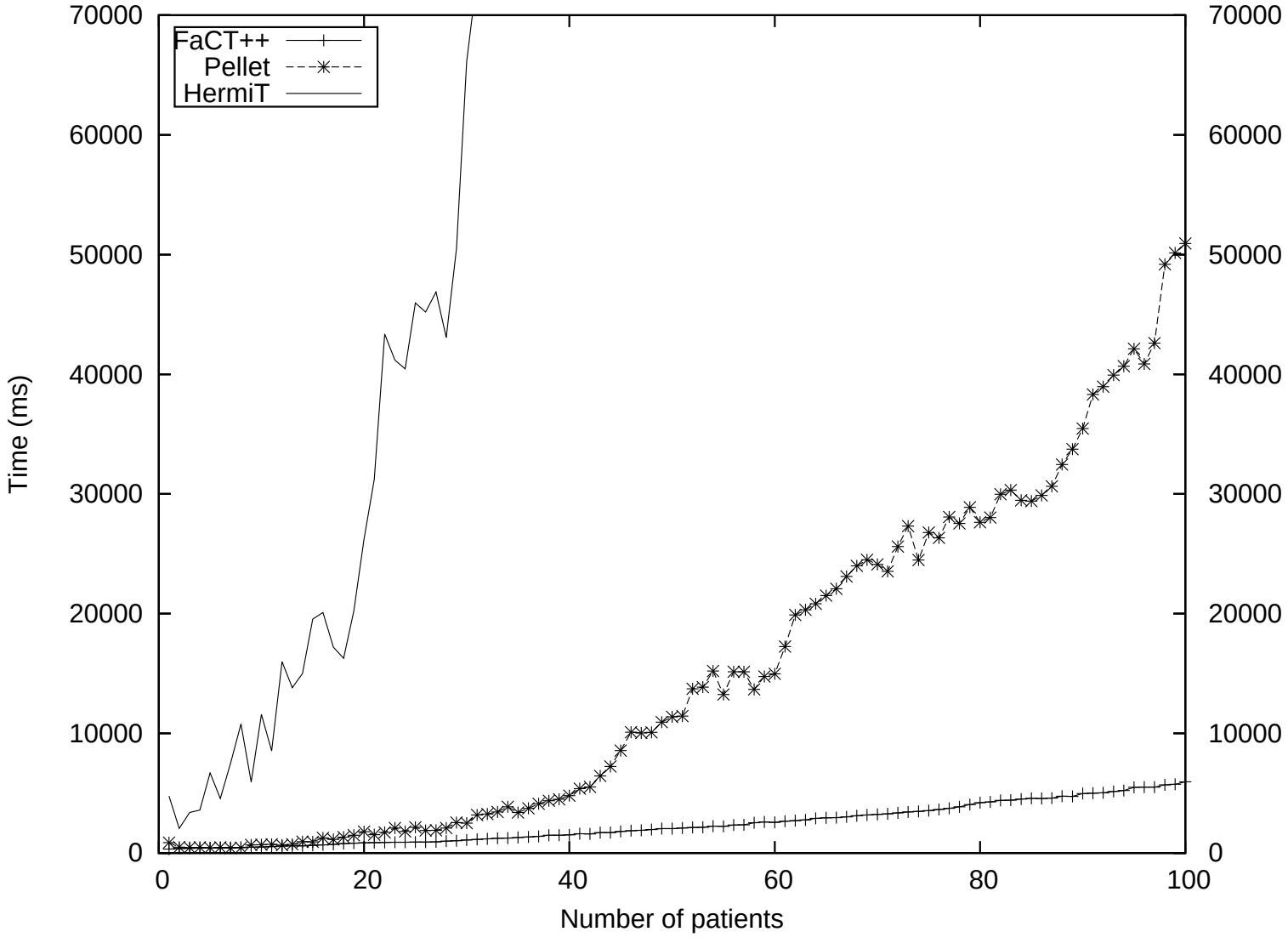


Evaluation

Experiment 2: Increasing the ABox with patient records

- 1 to 100 patient records.
- Experiment with LUCADA with 40 patient scenarios.
- Recorded times for realisation of all patients.

Experiment 2 (increasing ABox) with LUCADA



Conclusions from the LCA experiments

- **FaCT++ is currently the best choice for LCA.**
- HermiT provides the fastest TBox reasoning for LUCADA-SNOMED CT.
- HermiT does not scale for ABOX reasoning with LUCADA.
- Pellet performs well in classifying the LUCADA.
- Pellet struggles with the LUCADA-SNOMED CT ontology.

Questions?

- Lung Cancer Assistant (LCA):
  `http://lca.eng.ox.ac.uk/LungCancerSmartGWT/`
- LCA's main contact:
  Berkan Sesen (berkan.sesen@eng.ox.ac.uk)
- Tests and LUCADA-SNOMED integration:
  Ernesto Jimenez Ruiz (ernesto.jimenez.ruiz@gmail.com)

Thank you for your attention

Acknowledgements

- The LCA project was funded by the CDT in Healthcare Innovation programme within the Institute of Biomedical Engineering, Oxford University.