Queries, Modalities, Relations, Trees, XPath
Lecture I
Introduction

Tadeusz Litak

Department of Computer Science
University of Leicester

July 2010: draft

Recall the abstract . . .

Queries, Modalities, Relations, Trees, XPath

After some preliminaries on set theory and complexity theory, the first part of these lectures will contrast first-order logic (FOL), propositional modal logic (ML) and Tarski's relation algebra (TRA) as query languages for relational structures. Then we apply the general theory to finite node-labelled, sibling-ordered trees. Such structures represent XML trees as semantics for the navigational core of XPath — Core XPath 1.0 as introduced by Gottlob, Koch and Pichler and Core XPath 2.0 introduced by Marx and Ten Cate. Based on papers of the speaker, Balder ten Cate, Maarten Marx, Gaelle Fontaine and Carsten Lutz, we will show how to use existing ML techniques to characterize expressivity and complexity of XPath fragments and to provide complete axiomatizations of query equivalence and containment in these fragments.

The present one: a short overview of what we are going to do

Lecture II

Preliminaries: sets, relations, structures

An overview of basic set-theoretical notions, operations and constructions used in the lecture. Sets, ordered pairs and tuples, unions, intersections, products, powersets, relations as collections of ordered pairs, (graphs of) functions as special cases of relations, transitive closure, well-founded relations, labelled and unlabelled relational structures, trees and sibling-ordered trees as special cases of relational structures.

Slides are in the repository

Lecture III

Preliminaries: computability, complexity, models of computation

Lecture will be given on the whiteboard/blackboard, no slides in the repository. I will not go into details as this is not the main subject matter of the lecture anyway—we just need some understanding of notions involved.

Almost entire material will come from Appendix C of the following book:

Patrick Blackburn, Maarten de Rijke and Yde Venema.
Modal Logic
Cambridge Tracts in Theoretical Computer Science, 53,
Cambridge University Press, 2001.

Lecture III continued

The only addition might be a short discussion of models of computation different than Turing Machines, in particular

- 2-counter register machines
- Shepherdson-Sturgis register machines

This will be done for two reasons

- to illustrate that computations do not have to be modelled by Turing Machines
- such alternative models can be often easier to encode when we want to prove undecidability

Lecture IV

Correspondence Languages, Part 1

Definability and Expressivity

In this lecture, we will discuss various possible languages for labelled and unlabelled structures: First-Order Logic (FOL), Modal-Logic (ML), Tarski's Relation Algebras (TRAs), Dynamic Relation Algebras (DRAs) and Monadic Second-Order Logic (MSO), comparing and characterizing their expressive power.

Slides are in the repository

Lecture V

Correspondence Languages, Part 2

Axiomatization and Complexity

In this lecture we first introduce the notion of axiomatization and explain merits of having one. We will learn that increasing expressive power may come with a heavy price:

- computational complexity increases together with expressive power
- one can lose nice axiomatizations, like equational ones (contrast ML and DRA with TRA and FOL on infinite structures)
- when things go worse, one can lose decidability
- and in the worst case, one can lose not only decidability, but also any chance for axiomatization (MSO)

I have no slides for this lecture and will have to use whiteboard.

Harvest

This is a joint work with Balder ten Cate and Maarten Marx (University of Amsterdam). We provide complete axiomatizations for several fragments of Core XPath 1.0: the navigational core of XPath 1.0 introduced by Gottlob, Koch and Pichler. A complete axiomatization for a given fragment is a set of equivalences from which every other valid equivalence is derivable. Specifically, we axiomatize single axis fragments of Core XPath as well as full Core XPath 1.0. We also provide a short overview of complexity results for those fragments and a few others in between. A lot of attention is devoted to the issue of eliminating inelegant, "unorthodox" rewrite rules with side conditions.

Lecture VI, continued

. . .

This is a good opportunity to see a somewhat unexpected application of modal logic and algebra in CS, as the whole work is based on modal techniques for well-founded structures. In particular, idempotent couterdomain semirings (or antidomain semiring modules), the Loeb logic, "logic of finite trees" of Blackburn et al., logic of "dynamic negation" of Hollenberg, Fine's normal forms for modal formulas and reducts of Tarski's relation algebras are going to play a prominent role.

Slides of my last presentation on the subject in CVS/AIST, Osaka, November 2009 are in the repository.

Lecture VII (optional)

Still More Modal Aspects of XPath

In this talk, based mostly on a paper of B. ten Cate, G. Fontaine and myself, we would

- use general results discussed in the first part of this lecture to characterize expressivity of XPath fragments
- have a short look at axiomatization and decidability/complexity issues for fragments more expressive than Core XPath 1.0 (results of Marx, Ten Cate, Lutz . . .)

I would mostly use slides Balder ten Cate produced for his invited talk at M4M 2007. They are in the repository.

Lecture VIII (optional)

Termination problem for Declarative XML Message Processing

This lecture would present another application of classical results to modern WWW techologies and Web Services.

- We would define a simple Event-Condition-Action language
  for declarative XML message processing
- We would prove that

  for any practically useful fragment of this language
  the termination problem is undecidable
  — by encoding the Shepherdson-Sturgis register machine model of computation

The slides of my presentation at DEXA 2009 (based on a joint paper with Sven Helmer) are in the repository