Second-Order Quantifier Elimination on Relational Monadic Formulas – A Basic Method and Some Less Expected Applications

Christoph Wernhard

TU Dresden

TABLEAUX 2015

Wrozław, September 2015

Supported by DFG grant WE 5641/1-1

Second-Order Quantifier Elimination

- A computational task

  Input: A second-order formula

  $$\exists p\,(\forall x\, qx \rightarrow px) \wedge (\forall x\, px \rightarrow rx)$$

  Output: An equivalent first-order formula

  $$\forall x\, qx \rightarrow rx$$

- Related notions: uniform interpolation, forgetting, projection
- Applications
  - Computing frame correspondence properties of modal formulas
  - Computing circumscription and abductive explanations
  - Ontology reuse, ontology analysis, information hiding

Second-Order Quantifier Elimination: A Look Back into History

- Boole ∼1854, Peirce ∼1880, Schröder ∼1891: Algebra of Logic
- Löwenheim 1915, Skolem 1919: Deciding monadic formulas
- Behmann 1922
  Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem
  - Explicit formulation of the decision problem
  - Rewriting based decision/elimination method for monadic formulas
- Ackermann 1935
  Untersuchungen über das Eliminationsproblem der mathematischen Logik
  - Ackermann's Lemma, Ackermann's Quantifier Switching
  - Resolution-based elimination method
- Craig ∼1960
- Sahlqvist 1975, van Benthem 1983
  Computing modal first-order frame correspondence properties
- Gabbay/Ohlbach 1991
  Resolution-based elimination method (SCAN)
- Szałas/Doherty/Łukaszewicz 1992
  Elimination by rewriting to Ackermann's Lemma (DLS, Ackermann approach)
- Gabbay/Schmidt/Szałas 2008: Monograph
- Since ∼2006: Forgetting in description logics

Formula Classes Considered by Behmann

MON The class of relational monadic formulas (Löwenheim class)

i.e. first-order formulas
- with only nullary and unary predicates
- without functions except constants

$\mathrm{MON}_=$ MON with equality

QMON MON with predicate quantification

$\mathrm{QMON}_=$ MON with predicate quantification and equality

- All these classes are decidable
- Second-order quantifier elimination on a $\mathrm{QMON}_=$ formula always succeeds, resulting in a $\mathrm{MON}_=$ formula
- Elimination on a QMON formula might result in a $\mathrm{MON}_=$ formula

Behmann's Core Method and its Applications

- Behmann's core method eliminates $\exists p$ in $\exists p\, F$ where F is $\mathrm{MON}_=$
- It can be applied to eliminate second-order quantifiers in $\mathrm{QMON}_=$ formulas:
  - Repeatedly eliminate innermost second-order quantifiers
- It can be applied to decide $\mathrm{QMON}_=$ formulas:
  - Eliminate all second-order quantifiers in $\exists p_1 ... \exists p_n\ \exists x_1 ... \exists x_m\ \exists c_1 ... \exists c_k\, F$
  - The output is a $\mathrm{MON}_=$ sentence without predicates and constants

    It is either true or false for all domain cardinalities with exception of a finite number

    For all infinite domains the value is the same

Outline of Behmann's Elimination Method

Operates by equivalence preserving formula rewriting

1. Normalization to "innex form" w.r.t. first-order quantifiers
   - Propagate quantifiers inwards such that their scopes do not overlap
   - This is possible for MON and with counting quantifiers for $\mathrm{MON}_{=}$
   - Requires expensive rewritings such as distribution among $\land$ and $\lor$
   - Stronger than antiprenexing, e.g. [Egly, 1994]
2. Normalization such that subformulas $\exists p \ldots$ are in "Eliminationshauptform"
   - Involves potentially expensive inward propagation of $\exists p$
3. Conversion to match the left side of the Basic Elimination Lemma

Proposition: Basic Elimination Lemma

$$\exists p\, (\forall x\, Fx \lor px) \land (\forall x\, Gx \lor \neg px) \quad \equiv \quad \forall x\, Fx \lor Gx$$

4. Elimination by applying the Basic Elimination Lemma
5. Postprocessing

Other Decision Methods for Monadic Relational Formulas

- Innexing by a generalization of Shannon's expansion [Quine, 1945]
- Variants of resolution and superposition decide MON and $\mathrm{MON}_=$ [Fermüller et al., 2001, Bachmair et al., 1993]
- Complexity considerations [Lewis, 1980, Börger et al., 1997]
  - A satisfiable MON formula has a model whose cardinality is at most $2^{\text{number of predicates}}$
  - A model can be verified by a variant of innexing with DNF/CNF transformations where only atoms present in the input are involved

    The size never grows beyond a single exponential in the input size
  - Deciding satisfiability for MON and $\mathrm{MON}_=$ is NEXPTIME-complete

Elimination and Introduction of Auxiliary Definitions, Ackermann's Lemma

- It is common practice to obtain small equi-satisfiable CNFs by introducing auxiliary definitions – e.g. Tseitin's encoding
- Second-order quantification allows to understand the introduction and elimination of such definitions as equivalence preserving transformations

Proposition: Elimination and Introduction of Definitions

$$\exists p\, (\forall x\, px \leftrightarrow Gx) \land F[pt_1, \ldots, pt_n] \;\equiv\; F[Gt_1, \ldots, Gt_n]$$

Proposition: Ackermann's Lemma [Ackermann, 1935a]

If all n indicated subformula occurrences in $F[pt_1, \ldots, pt_n]$:
(or, equivalently, in $F[Gt_1, \ldots, Gt_n]$) have the same polarity POL, then

$$\exists p\, (\forall x\, px \rightarrow Gx) \land F[pt_1, \ldots, pt_n] \;\equiv\; F[Gt_1, \ldots, Gt_n], \text{ if } POL = \text{pos}$$
$$\exists p\, (\forall x\, px \leftarrow Gx) \land F[pt_1, \ldots, pt_n] \;\equiv\; F[Gt_1, \ldots, Gt_n], \text{ if } POL = \text{neg}$$

Ackermann's Quantifier Switching

Proposition: Ackermann's Quantifier Switching [Ackermann, 1935b]

$$\exists p\, \forall x\, F[pxt_1, \ldots, pxt_n] \;\equiv\; \forall x\, \exists p'\, F[p't_1, \ldots, p't_n]$$

- [Ackermann, 1935b] applies it to avoid Skolemization and to allow monadic techniques or Ackermann's Lemma
- [van Benthem, 1983] applies it right-to-left to achieve prenex form w.r.t. second-order quantifiers

Normalized Representation of Description Logic KBs in Classical Logics

- A description logic (DL) knowledge base (KB) is a pair $\langle \text{TBox}, \text{ABox} \rangle$
- Many DLs can be considered as fragments of first-order logic
- A KB in such a DL can be translated to a first-order sentence, based on the standard translation of modal logics
- This representation can be normalized to a second-order formula where
  - Quantified auxiliary concept predicates "define" subformulas
  - The matrix is a generalized CNF, in place of literals are basic formulas:

Basic formula	*Inducing DL construct*
cx	atomic concept, ABox assertion
$\neg cx$	atomic concept
$\exists y\, rxy \land dy$	qualified existential restriction
$\forall y\, \neg rxy \lor dy$	qualified value restriction
$x \neq a$	ABox assertion
rxa	ABox assertion

  - Observe: All occurrences of role predicates have x as first argument

Normalized Representation of Description Logic KBs: Example

$$
\begin{array}{rcrcll}
KB & = & \langle \{ C_1 & \sqsubseteq & \exists R.(C_2 \sqcup C_3), & \\
 & & C_2 & \sqsubseteq & C_4 \}, & \\
 & & \{ C_1(A) \} \rangle & & & \\
\\
\forall x\, \mathrm{ST}_x(KB) & = & \forall x \quad c_1 x & \rightarrow & \exists y\, r x y \land (c_2 y \lor c_3 y) & \land \\
 & & c_2 x & \rightarrow & c_4 x & \land \\
 & & c_1 a & & & \\
\\
\exists d_1\, \forall x\, F & = & \exists d_1\, \forall x \quad c_1 x & \rightarrow & \exists y\, r x y \land d_1 y & \land \\
 & & d_1 x & \rightarrow & c_2 x \lor c_3 x & \land \\
 & & c_2 x & \rightarrow & c_4 x & \land \\
 & & x = a & \rightarrow & c_1 x &
\end{array}
$$

Satisfiability of $\mathcal{ALC}$ as $\mathrm{QMON}_=$ Satisfiability

- The normalized translation

$$\exists d_1 \ldots \exists d_k \, \forall x \, F$$

is equi-satisfiable with the closure of its free concept and role predicates:

$$\exists c_1 \ldots \exists c_n \, \exists r_1 \ldots \exists r_m \, \exists d_1 \ldots \exists d_k \, \forall x \, F$$

- The predicate quantifiers can be reordered:

$$\exists c_1 \ldots \exists c_n \, \exists d_1 \ldots \exists d_k \, \exists r_1 \ldots \exists r_m \, \forall x \, F$$

- Ackermann's Quantifier Switching can be applied to the role predicates:

$$\exists c_1 \ldots \exists c_n \, \exists d_1 \ldots \exists d_k \, \forall x \, \exists r'_1 \ldots \exists r'_m \, F[r_i x t \mapsto r'_i t]$$

- This is a $\mathrm{QMON}_=$ formula
- This technique applies to $\mathcal{ALCOQH}$: atomic concepts, $\top$, $\bot$, $\neg$, $\sqcap$, $\exists R.C$, $\forall R.C$ + nominals, qualified number restrictions, subroles

Satisfiability of $\mathcal{ALCOQH}$ KBs can be expressed as $\mathrm{QMON}_=$ satisfiability

Elimination for DLs as $\mathbf{QMON}_=$ Elimination (I)

- The objective is to apply second-order quantifier elimination to

$$\exists p_1 \ldots \exists p_n \, \forall x \, \mathbf{ST}_x(KB),$$

  where the p_i are concept predicates

- Normalization and further conversion yields an equivalent formula

$$S \land \exists c_1 \ldots \exists c_l \, \exists d_1 \ldots \exists d_k \, \forall x \, F,$$

  where

  - S is a sentence with only role predicates
  - c_i are concept predicates to eliminate that only occur with argument x:
    - the p_i
    - some of the auxiliary concept predicates
  - d_i are the remaining auxiliary concept predicates
  - F is a generalized CNF

Elimination for DLs as $\mathrm{QMON}_=$ Elimination (II)

- We introduce unary role predicates r'_i and store the original r_i in definitions:

  Let Rx be $\bigwedge_{1 \leq i \leq m} (\forall y\, r'_i y \leftrightarrow r_i x y)$. Then our formula

  $$S \wedge \exists c_1 \ldots \exists c_l\, \exists d_1 \ldots \exists d_k\, \forall x\, F$$

  is equivalent to

  $$S \wedge \exists c_1 \ldots \exists c_l\, \exists d_1 \ldots \exists d_k\, \forall x\, \exists r'_1 \ldots \exists r'_m\, Rx \wedge F[r_i x t \mapsto r'_i t]$$

- We apply Ackermann's Quantifier Switching to the c_i:

  $$S \wedge \exists d_1 \ldots \exists d_k\, \forall x\, \exists r'_1 \ldots \exists r'_m\, Rx \wedge \exists c'_1 \ldots \exists c'_l\, F[r_i x t \mapsto r'_i t,\, c_i x \mapsto c'_i]$$

- $F[\ldots]$ is now a $\mathrm{MON}_=$ formula, thus we can eliminate the $\exists c'_i$
- We then restore the binary role predicates according to Rx and delete Rx
- Our formula then has the form

  $$S \wedge \exists d_1 \ldots \exists d_k\, \forall x\, F'$$

  and we are finished if there are no d_i
- This is the case if the DL has only limited (unqualified) restriction
  e.g. $\exists R.\top$

Elimination for DLs as $\mathbf{QMON}_=$ Elimination (III)

Concept elimination on KBs in $\mathcal{ALCOQH}$ with only limited restriction but in addition inverse roles can be performed by second-order quantifier elimination on $\mathbf{QMON}_=$ and definition elimination

- With inverse roles but only limited restriction, this covers typical members of the DL-Lite family

Related Approaches in Description Logics

- Two-phase forgetting for expressive DLs
  [Koopmann and Schmidt, 2013b, Koopmann and Schmidt, 2015]
  - Based on resolution and a generalized Ackermann's Lemma
  - The phases are roughly related to the c_i and d_i
  - Normalization is preserved throughout to allow re-translation to DL
- Equi-satisfiable translations of DL-Lite into the one-variable fragment
  [Artale et al., 2009]
- Specialized methods for forgetting in DL-Lite
  [Kontchakov et al., 2010, Wang et al., 2010]

The Ackermann Approach to Second-Order Quantifier Elimination

- Also called direct approach
- Basic idea: Rewrite such that elimination can be performed by applying Ackermann's Lemma to subformulas
- Exemplified by DLS
  [Szałas, 1993, Doherty et al., 1997, Gustafsson, 1996, Conradie, 2006]
  1. Preprocessing to a certain normal form
     This might fail
  2. Preparation for Ackermann's Lemma
     Skolem functions might be introduced
  3. Application of Ackermann's Lemma and un-Skolemization
     Un-Skolemization might fail
  4. Simplification

A Limitation of DLS – Obvious in View of Behmann's Method

- The preprocessing of DLS fails on some monadic inputs that, however, could be converted to the target normal form

  The reason is that DLS first converts to negation normal form and then includes distribution of $\wedge$ over $\vee$, but not distribution of $\vee$ over $\wedge$

Ackermann's Lemma for "Semi Monadic" Formulas

2. Preparation for Ackermann's Lemma

   Convert $\exists p\, A \land B$, where A (B) contains p just positively (negatively), to

   $$\exists f_1 \ldots \exists f_k\, \exists p\, (\forall x\, A'x \rightarrow px) \land B'$$

3. Application of Ackermann's Lemma and un-Skolemization

   Un-Skolemization might fail

If A or B is a $\mathbf{MON}_=$ formula, the preparation for Ackermann's Lemma can be performed without introduction of Skolem functions

- The result may contain several subformulas $\exists p \ldots$ that all match Ackermann's Lemma
- Elimination on Sahlqvist formulas matches the "semi monadic" case
  - Success on computing correspondence properties of Sahlqvist formulas is an investigated completeness property of elimination methods [Goranko et al., 2004, Conradie, 2006, Schmidt, 2012]
  - The core step of a special elimination method for Sahlqvist formulas (the Sahlqvist-van Benthem substitution algorithm) can be considered as application of Ackermann's Lemma on a "semi monadic" input

Open and Future Issues

- **Further investigation of methods for relational monadic formulas**
  - Complexity of QMON, $\mathrm{QMON}_=$?
  - How can particular formula patterns occurring in applications be utilized?
  - What implementation techniques are relevant?
- **Further investigation of applications**
  - Taking finite relations and constraint databases into account
  - Deeper investigation of forgetting in DLs
  - First-order consequences vs. DL consequences
- **Follow-up issues on elimination**
  - Further completeness criteria of elimination methods
  - Further improvements of DLS
  - Investigation of Ackermann's resolution-based method and related works

Summary

- **Restoration of Behmann's second-order quantifier elimination method**
  - Continued work by Schröder, inspired Ackermann
  - Developed along the notion of Entscheidungsproblem
  - Explicitly targeted at practical application – Computational Logic
  - Operates by equivalence preserving formula rewriting to normal forms
- **Observing aspects of Behmann's method that seem not yet digested**
  - Similarity to modern methods of the Ackermann approach
  - Beyond antiprenexing: innexing for the price of expensive operations
  - Completeness for monadic formulas as criterion for elimination methods
- **Expressing applications as elimination on relational monadic formulas**
  - Deciding $\mathcal{ALC}$ knowledge bases
  - Concept elimination in DL-Lite knowledge bases
  - Eliminability in "semi monadic" formulas, for example Sahlqvist formulas
- **Use of a methodology with equivalence preserving transformations**
  - First-order based, applying second-order properties like Ackermann's Quantifier Switching and Ackermann's Lemma

References

[Ackermann, 1935a] Ackermann, W. (1935a).
Untersuchungen über das Eliminationsproblem der mathematischen Logik.
Mathematische Annalen, 110:390–413.

[Ackermann, 1935b] Ackermann, W. (1935b).
Zum Eliminationsproblem der mathematischen Logik.
Mathematische Annalen, 111:61–63.

[Artale et al., 2009] Artale, A., Calvanese, D., Kontchakov, R., and Zakharyaschev, M. (2009).
The DL-Lite family and relations.
JAIR, (36):1–69.

[Bachmair et al., 1993] Bachmair, L., Ganzinger, H., and Waldmann, U. (1993).
Superposition with simplification as a decision procedure for the monadic class with equality.
In *KGC'93*, volume 713 of *LNCS*, pages 83–96.

[Behmann, 1922] Behmann, H. (1922).
Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem.
Mathematische Annalen, 86(3–4):163–229.

[Börger et al., 1997] Börger, E., Grädel, E., and Gurevich, Y. (1997).
The Classical Decision Problem.
Springer.

[Conradie, 2006] Conradie, W. (2006).
On the strength and scope of DLS.
JANCL, 16(3–4):279–296.

[Craig, 1960] Craig, W. (1960).
Bases for first-order theories and subtheories.
JSL, 25(2):97–142.

[Craig, 2008] Craig, W. (2008).
Elimination problems in logic: A brief history.
Synthese, (164):321–332.

[Doherty et al., 1997] Doherty, P., Łukaszewicz, W., and Szałas, A. (1997).
Computing circumscription revisited: A reduction algorithm.
JAR, 18(3):297–338.

[Egly, 1994] Egly, U. (1994).
On the value of antiprenexing.
In *LPAR '94*, volume 822 of *LNCS*, pages 69–83.

[Fermüller et al., 2001] Fermüller, C., Leitsch, A., Hustadt, U., and Tammet, T. (2001).
Resolution decision procedures.
In Robinson, A. and Voronkov, A., editors, *Handbook of Automated Reasoning*, volume 2, pages 1793–1849. Elsevier.

[Gabbay et al., 2008] Gabbay, D. M., Schmidt, R. A., and Szałas, A. (2008).
Second-Order Quantifier Elimination: Foundations, Computational Aspects and Applications.
College Publications.

[Goranko et al., 2004] Goranko, V., Hustadt, U., Schmidt, R. A., and Vakarelov, D. (2004).
SCAN is complete for all Sahlqvist formulae.
In *RelMiCS 7*, volume 3051 of *LNCS*, pages 149–162.

[Gustafsson, 1996] Gustafsson, J. (1996).
An implementation and optimization of an algorithm for reducing formulae in second-order logic.
Technical Report LiTH-MAT-R-96-04, Univ. Linköping.

[Kontchakov et al., 2010] Kontchakov, R., Wolter, F., and Zakharyaschev, M. (2010).
Logic-based ontology comparison and module extraction, with an application to DL-Lite.
AI, 174(15):1093–1141.

[Koopmann and Schmidt, 2013a] Koopmann, P. and Schmidt, R. A. (2013a).
Forgetting concept and role symbols in $\mathcal{ALCH}$-ontologies.
In *LPAR-19*, volume 8312 of *LNCS (LNAI)*, pages 552–567. Springer.

[Koopmann and Schmidt, 2013b] Koopmann, P. and Schmidt, R. A. (2013b).
Uniform interpolation of $\mathcal{ALC}$-ontologies using fixpoints.
In *FroCoS 2013*, volume 8152 of *LNCS (LNAI)*, pages 87–102. Springer.

[Koopmann and Schmidt, 2015] Koopmann, P. and Schmidt, R. A. (2015).
Uniform interpolation and forgetting for $\mathcal{ALC}$ ontologies with ABoxes.
In *AAAI 15*, pages 175–181. AAAI Press.

[Lewis, 1980] Lewis, H. R. (1980).
Complexity results for classes of quantificational formulas.
J. Computer and System Sciences, 21:317–353.

[Löwenheim, 1915] Löwenheim, L. (1915).
Über Möglichkeiten im Relativkalkül.
Mathematische Annalen, 76:447–470.

[Quine, 1945] Quine, W. V. (1945).
On the logic of quantification.
JSL, 10(1):1–12.

[Sahlqvist, 1975] Sahlqvist, H. (1975).
Completeness and correspondence in the first and second order semantics for modal logic.
In *Proc. Third Scand. Logic Symp. Uppsala (1973)*, pages 110–143. North-Holland.

[Schmidt, 2012] Schmidt, R. A. (2012).
The Ackermann approach for modal logic, correspondence theory and second-order reduction.
JAL, 10(1):52–74.

[Schröder, 1905] Schröder, E. (1890–1905).
Vorlesungen über die Algebra der Logik.
Teubner.

[Skolem, 1919] Skolem, T. (1919).
Untersuchungen über die Axiome des Klassenkalküls und über Produktations- und Summationsprobleme welche gewisse Klassen von Aussagen betreffen.
Videnskapsselskapets Skrifter, I. Mat.-Nat. Klasse(3).

[Szałas, 1993] Szałas, A. (1993).
On the correspondence between modal and classical logic: An automated approach.
JLC, 3(6):605–620.

[van Benthem, 1983] van Benthem, J. (1983).
Modal Logic and Classical Logic.
Bibliopolis, Napoli.

[Wang et al., 2010] Wang, Z., Wang, K., Topor, R. W., and Pan, J. Z. (2010).
Forgetting for knowledge bases in DL-Lite.
Annals of Mathematics and Artificial Intelligence, 58:117–151.