Towards More Effective Tableaux Reasoning for CKR

1Loris Bozzato  \hspace{1cm}  2Martin Homola  \hspace{1cm}  1Luciano Serafini

1DKM, Fondazione Bruno Kessler – Trento, Italy
2Comenius University – Bratislava, Slovakia

25th Int. Workshop on Description Logics (DL 2012)
June 7-10, 2012 - Roma, Italy
1. Introduction
2. Contextualized Knowledge Repositories
3. Tableaux Algorithm
4. Algorithm Optimization
5. Conclusions
Motivation

Need for context in Semantic Web

- Most of Semantic Web data holds in specific **contextual space** (time, location, topic...)
- **No explicit support** for reasoning with context sensitive knowledge in SW (Often handcrafted in implementation)
- Need for well-defined theory of contexts
Motivation

Need for context in Semantic Web

- Most of Semantic Web data holds in specific **contextual space** (time, location, topic...)
- **No explicit support** for reasoning with context sensitive knowledge in SW (Often handcrafted in implementation)
  
  ➔ Need for well-defined theory of contexts

Contextualized Knowledge repository (CKR)

- DL based framework for representation and reasoning with contextual knowledge in the Semantic Web
  
  - **Contextual theory**: based on formal AI theories of context

Other DL contextual frameworks:

[Bao et al., 2010, Klarman and Gutiérrez-Basulto, 2011, Straccia et al., 2010].
“Context as a Box” paradigm

Idea [Benerecetti et al., 2000]

- A context is a logical theory...
- ...associated to a region in a contextual space
“Context as a Box” paradigm

Idea [Benerecetti et al., 2000]

- A context is a logical theory...
- ...associated to a region in a contextual space

\[
C = \text{HostTeam} \sqsubseteq \text{QualifiedTeam}
\]

\[
\ldots
\]

\[
\text{Winner(team_spain)}
\]

\[
\text{RunnerUp(team_holland)}
\]

\[
\ldots
\]

\[
\text{playsFor(buffon, team_italy)}
\]

\[
\text{playsFor(cannavaro, team_italy)}
\]

\[
\ldots
\]
“Context as a Box” paradigm

Idea [Benerecetti et al., 2000]

- A context is a logical theory...
- ...associated to a region in a contextual space

\[
C = \begin{cases} 
\text{HostTeam} \sqsubseteq \text{QualifiedTeam} \\
\quad \ldots \\
\quad \text{Winner(team\_spain)} \\
\quad \text{RunnerUp(team\_holland)} \\
\quad \ldots \\
\quad \text{playsFor(buffon, team\_italy)} \\
\quad \text{playsFor(cannavaro, team\_italy)} \\
\quad \ldots 
\end{cases}
\]

\[
\text{time}(C, 2010), \ \text{location}(C, \text{South\_Africa}), \ \text{topic}(C, \text{FIFA\_WC})
\]
“Context as a Box” paradigm

Idea [Benerecetti et al., 2000]

- A context is a **logical theory**...
- ...associated to a region in a **contextual space**

![Diagram of context as a box with dimensions for space, time, and topic. The box contains information about the FIFA World Cup 2010 in South Africa with winners and runners-up.]
Contextual dimensions [Lenat, 1998]

- time, location, topic, … are dimensions
- Our logical framework supports $n$ dimensions
- Each dimension $A$ ordered w.r.t. coverage relation $\prec_A$
Dimensions and coverage

Contextual dimensions [Lenat, 1998]
- time, location, topic, ... are dimensions
- Our logical framework supports $n$ dimensions
- Each dimension $A$ ordered w.r.t. coverage relation $\prec_A$

Context coverage
- Narrower / broader context relation defined from dimension ordering
- This relation defines the covering structure of contexts
Hierarchy of contexts

< 2010, World, Sports >

< 2010, World, Football >

< 2010, S.Africa, FIFA_WC >

< 2010, Italy, National_League >
Reference to other contexts

- **Locality of knowledge**
  - Concept $C$ and role $R$ symbols have independent meaning in different contexts

- **Qualified symbols**: refer to entities from other contexts
  - For every concept $C$ and role $R$
  - and every dimensional vector $d$
  - we introduce qualified symbols $C_d$ and $R_d$

$$C = \text{time}(C, 2010), \text{location}(C, \text{South_Africa}), \text{topic}(C, \text{FIFA_WC})$$

\[
\ldots \\
\text{playsFor}(\text{buffon, team_italy}) \\
\text{playsFor}_{\text{World, Italian_league}}(\text{buffon, juventus}) \\
\text{playsFor}(\text{alcaraz, team_paraguay}) \\
\text{playsFor}_{\text{World, English_league}}(\text{alcaraz, wigan}) \\
\ldots
\]
Outline

1. Introduction
2. Contextualized Knowledge Repositories
3. Tableaux Algorithm
4. Algorithm Optimization
5. Conclusions
Meta and Object vocabulary

**Meta vocabulary \( \Gamma \)**

- **dimensions** \( A \in A \)
  - roles: topic, time, location, . . .
- **dimensional values** \( a \in D_A \)
  - individuals: FIFAWC, football, . . .
- **coverage relation** \( A \)
  - roles: FIFAWC \( \prec_{\text{topic}} \) football, . . .
Meta and Object vocabulary

Meta vocabulary $\Gamma$

- **dimensions** $A \in A$
  - roles: topic, time, location, . . .
- **dimensional values** $a \in D_A$
  - individuals: FIFAWC, football, . . .
- **coverage relation** $\prec_A$
  - roles: FIFAWC $\prec_{\text{topic}}$ football, . . .

Dimensional vectors $d = \{d_1, \ldots, d_m\}$ with $d_i \in A_i$

- e.g. $\{2010, \text{world, football}\}$
- Partial or full
- Operations: $d \prec e, d + e$
### Meta vocabulary $\Gamma$

- **dimensions** $A \in A$
  - roles: topic, time, location, ...
- **dimensional values** $a \in D_A$
  - individuals: FIFAWC, football, ...
- **coverage relation** $\prec_A$
  - roles: FIFAWC $\prec_\text{topic}$ football, ...

**Dimensional vectors** $\mathbf{d} = \{d_1, \ldots, d_m\}$ with $d_i \in A_i$

- e.g. $\{2010, \text{world}, \text{football}\}$
- Partial or full
- Operations: $\mathbf{d} \prec \mathbf{e}, \mathbf{d} + \mathbf{e}$

### Object vocabulary $\Sigma$

DL vocabulary closed by **qualification** on concepts ($C_d$) and roles ($R_d$)
CKR definition

- **CKR knowledge base** $\mathcal{K} = \langle \mathcal{M}, \mathcal{C} \rangle$ on $(\Gamma, \Sigma)$

- **$\mathcal{C}$**: finite collection of contexts $C_d$
  - $\text{dim}(C_d) = d$, a full dimensional vector of $\Gamma$
  - $K(C_d)$, an $\text{ALC}$ KB over $\Sigma$

- **Meta knowledge** $\mathcal{M}$: DL KB with meta assertions
  - Dimensional values: $D_A = \{d, e, \ldots\}$
  - Dimensional coverage: $d \prec_A e$
  - Context dimensions relations: $A(C, d)$
CKR semantics

Interpretation for CKR knowledge bases:
- One DL interpretation for each context
- Semantic conditions for knowledge compatibility

Partial DL interpretation

DL interpretation \( \mathcal{I} = \langle \Delta^\mathcal{I}, .^\mathcal{I} \rangle \) in which:
- \( \Delta^\mathcal{I} \) can be empty
- \( .^\mathcal{I} \) can be partially defined on individuals
**CKR Model**

**Set of partial interpretations:** \( \mathcal{I} = \{ \mathcal{I}_d \}_{d \in \mathcal{D}} \)

**Rules for propagation and locality of knowledge:**

1. \( (\top_d)^{\mathcal{I}_f} \subseteq (\top_e)^{\mathcal{I}_f} \) if \( d \prec e \) \hspace{1cm} (domain containment)
2. \( (A_f)^{\mathcal{I}_d} \subseteq (\top_f)^{\mathcal{I}_d} \) \hspace{1cm} (local interpretation 1)
3. \( (R_f)^{\mathcal{I}_d} \subseteq (\top_f)^{\mathcal{I}_d} \times (\top_f)^{\mathcal{I}_d} \) \hspace{1cm} (local interpretation 2)
4. \( a^{\mathcal{I}_e} = a^{\mathcal{I}_d} \) if \( a^{\mathcal{I}_e} \in \Delta_d \) and \( d \prec e \) or \( d \succ e \) \hspace{1cm} (global names)
5. \( (X_{d_B})^{\mathcal{I}_e} = (X_{d_B} + e)^{\mathcal{I}_e} \) \hspace{1cm} (partial qualification)
6. \( (X_d)^{\mathcal{I}_e} = (X_d)^{\mathcal{I}_d} \) if \( d \prec e \) \hspace{1cm} (restricted interpretation 1)
7. \( (A_f)^{\mathcal{I}_d} = (A_f)^{\mathcal{I}_e} \cap \Delta_d \) if \( d \prec e \) \hspace{1cm} (restricted interpretation 2)
8. \( (R_f)^{\mathcal{I}_d} = (R_f)^{\mathcal{I}_e} \cap (\Delta_d \times \Delta_d) \) if \( d \prec e \) \hspace{1cm} (restricted interpretation 3)
9. \( \mathcal{I}_d \models K(C_d) \)
CKR Model $\mathcal{I}$

- **Set of partial interpretations:** $\mathcal{I} = \{\mathcal{I}_d\}_{d \in \mathbb{D}}$

- **Rules for propagation and locality of knowledge:**
  1. $(\top_d)^{\mathcal{I}_f} \subseteq (\top_e)^{\mathcal{I}_f}$ if $d \prec e$ (domain containment)
  2. $(A_f)^{\mathcal{I}_d} \subseteq (\top_f)^{\mathcal{I}_d}$ (local interpretation 1)
  3. $(R_f)^{\mathcal{I}_d} \subseteq (\top_f)^{\mathcal{I}_d} \times (\top_f)^{\mathcal{I}_d}$ (local interpretation 2)
  4. $a^{\mathcal{I}_e} = a^{\mathcal{I}_d}$ if $a^{\mathcal{I}_e} \in \Delta_d$ and $d \prec e$ or $d \succ e$ (global names)
  5. $(X_{d_B})^{\mathcal{I}_e} = (X_{d_B + e})^{\mathcal{I}_e}$ (partial qualification)
  6. $(X_d)^{\mathcal{I}_e} = (X_d)^{\mathcal{I}_d}$ if $d \prec e$ (restricted interpretation 1)
  7. $(A_f)^{\mathcal{I}_d} = (A_f)^{\mathcal{I}_e} \cap \Delta_d$ if $d \prec e$ (restricted interpretation 2)
  8. $(R_f)^{\mathcal{I}_d} = (R_f)^{\mathcal{I}_e} \cap (\Delta_d \times \Delta_d)$ if $d \prec e$ (restricted interpretation 3)
  9. $\mathcal{I}_d \models K(C_d)$
**d-satisfiability**

*C is d-satisfiable w.r.t. \( \mathcal{K} \) if there exists \( \mathcal{I} = \{ \mathcal{I}_e \}_{e \in D} \) of \( \mathcal{K} \) s.t. \( C^{\mathcal{I}_d} \neq \emptyset \).*

**d-entailment**

*\( \alpha \) is d-entailed by \( \mathcal{K} \) if for every model \( \mathcal{I} = \{ \mathcal{I}_e \}_{e \in D} \) of \( \mathcal{K} \), \( \mathcal{I}_d \models \alpha \).*
Tableaux algorithm $C_T$

Tableaux algorithm $C_T$ for $\mathcal{ALC}$-based CKR

- Decides $\mathfrak{d}$-satisfiability of concept $C$ w.r.t. $\mathcal{K}$
- Extension of $\mathcal{ALC}$ non-deterministic algorithm
- Single completion tree, but different labels for each context

Tableaux rules:

- local reasoning: usual $\mathcal{ALC}$ rules
- knowledge propagation: new CKR rules
Completion tree

Completion tree:
Partial representation of a CKR model

- $T = \langle V, E, \mathcal{L} \rangle$ s.t. for each $e$:
  - nodes: $V_e \subseteq V$
  - edges: $E_e \subseteq E$ and $E_e \subseteq V_e \times V_e$
  - labels: $\mathcal{L} = \{ \mathcal{L}_e \}_{e \in \mathfrak{D}_T}$ with:
    - $\mathcal{L}_e(x)$ set of concepts for $x \in V_e$
    - $\mathcal{L}_e(\langle x, y \rangle)$ set of roles for $\langle x, y \rangle \in E_e$
Algorithm definition

Initialization: \( T := \langle V, E, L \rangle \)

- \( V_e := \{a^e | C(a) \in K(C_e)\} \cup \{a^e, b^e | R(a, b) \in K(C_e)\} \)
- \( E_e := \{\langle a^e, b^e \rangle | R(a, b) \in K(C_e)\} \)
- \( \mathcal{L}_e(a^e) := \{C | C(a) \in K(C_e)\} \)
- \( \mathcal{L}_e(\langle a^e, b^e \rangle) := \{R | R(a, b) \in K(C_e)\} \)
- \( V_d := V_d \cup \{s_0\}, \text{ with } s_0 \text{ new} \)
- \( \mathcal{L}_d(s_0) := \{C\} \)

Tableaux expansion:

- apply all rules exhaustively
- try to construct complete and clash-free c.tree

Note

- **Blocking:** equality blocking on all \( \mathcal{L}_e \) [Buchheit et al., 1993]
- **Merging:** \( M \)-rule merges ABox individuals with same meaning
Tableaux rules: \( \mathcal{ALC} \) local rules

<table>
<thead>
<tr>
<th>( \mathcal{L}_d(x) = {C_1 \sqcap C_2} )</th>
<th>( \Box )</th>
<th>( \mathcal{L}_d(x) = {C_1 \sqcap C_2, C_1, C_2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{L}_d(x) = {C_1 \sqcup C_2} )</td>
<td>( \top )</td>
<td>( \mathcal{L}_d(x) = {C_1 \sqcup C_2, C} ) for ( C \in {C_1, C_2} )</td>
</tr>
<tr>
<td>( \bullet \mathcal{L}_d(x) = {\exists R.C} )</td>
<td>( \exists )</td>
<td>( \bullet \mathcal{L}_d(x) = {\exists R.C} )</td>
</tr>
<tr>
<td>( \bullet \mathcal{L}_d(x) = {\forall R.C} ) ( \mathcal{L}_d(&lt;x,y&gt;) = {R} )</td>
<td>( \forall )</td>
<td>( \bullet \mathcal{L}_d(x) = {\forall R.C} ) ( \mathcal{L}_d(&lt;x,y&gt;) = {R} )</td>
</tr>
<tr>
<td>( \mathcal{L}_d(x) = {\ldots}, {C \sqsubseteq D} \subseteq K(\mathcal{C}_d) )</td>
<td>( \mathcal{T} )</td>
<td>( \mathcal{L}_d(x) = {\text{nnf}(\neg C \sqcup D)} )</td>
</tr>
</tbody>
</table>
Tableaux rules: CKR propagation rules

\[ \begin{array}{c|c|c}
\text{d} & \Delta \uparrow & \Delta \downarrow \\
\hline
\text{x} & \bullet & \bullet \\
\text{L}_d(x) = \{ \top_d \} & \bullet & \bullet \\
\text{L}_e(x) = \{ A_f \} & \bullet & \bullet \\
\hline
\text{R} & \text{R} & \text{R} \\
\text{x} & \bullet & \bullet \\
\text{e} & \bullet & \bullet \\
\end{array} \]
Tableaux rules: CKR top and M-rules

<table>
<thead>
<tr>
<th>$\mathcal{L}_d(x) = {A_f}$</th>
<th>$\top_A$</th>
<th>$\mathcal{L}_d(x) = {A_f, \top_f}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bullet \mathcal{L}_d(x) = {\ldots}$</td>
<td>$\top_R$</td>
<td>$\bullet \mathcal{L}_d(x) = {\top_f}$</td>
</tr>
<tr>
<td>$\mathcal{L}_d(&lt;x,y&gt;) = {R_f}$</td>
<td></td>
<td>$\mathcal{L}_d(&lt;x,y&gt;) = {R_f}$</td>
</tr>
<tr>
<td>$\bullet \mathcal{L}_d(y) = {\ldots}$</td>
<td></td>
<td>$\bullet \mathcal{L}_d(y) = {\top_f}$</td>
</tr>
</tbody>
</table>

| $\mathcal{L}_d(x) = \{\ldots\}, e \preccurlyeq f$ | $\subseteq$ | $\mathcal{L}_d(x) = \{\neg \top_e \cup \top_f\}, e \ll f$ |

<table>
<thead>
<tr>
<th>$d$</th>
<th>$\bullet a^h$</th>
<th>$e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bullet a^g$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
</tr>
<tr>
<td>$\bullet a$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$d$</th>
<th>$\bullet a$</th>
<th>$e$</th>
</tr>
</thead>
</table>
\[ \mathcal{A} \models d : A_e \sqsubseteq B_f \]
Deduction example II

\[ A \equiv C_f \]

\[ \forall \mathcal{R} \models d : A_e \sqsubseteq B_f \]

\[ s_0 : A_e \sqcap \neg B_f \]

\[ C \sqsubseteq B \]
\( A \equiv C_f \)

\( C \sqsubseteq B \)

\( s_0 : A_e \sqcap \neg B_f, \quad A_e, \neg B_f \)

\( \sqcap \)-rule

\( \mathcal{K} \models d : A_e \sqsubseteq B_f \)
Deduction example IV

\[ A \equiv C_f, \]
\[ s_0 : A \]

\[ C \sqsubseteq B \]

\[ s_0 : A_e \sqcap \neg B_f, \]
\[ A_e, \neg B_f \]

\[ \Delta \uparrow, A\text{-rules} \]

\[ \mathcal{H} \models d : A_e \sqsubseteq B_f \]
Deduction example V

\[ A \equiv C_f, \]
\[ s_0 : A, C_f \]

\[ C \sqsubseteq B \]
\[ s_0 : A_e \sqcap \neg B_f, \]
\[ A_e, \neg B_f \]

\[ \mathcal{T}, \sqcap\text{-rules} \]
\[ \mathcal{R} \models d : A_e \sqsubseteq B_f \]
Deduction example VI

\[ A \equiv C_f, \]
\[ s_0 : A, C_f, T_f \]

\[ C \sqsubseteq B, \]
\[ s_0 : C \]

\[ s_0 : A_e \sqcap \neg B_f, \]
\[ A_e, \neg B_f \]

\[ \top, \Delta \downarrow, A\text{-rules} \]

\[ \mathcal{K} \models d : A_e \sqsubseteq B_f \]
\[
A \equiv C_f, \\
\mathcal{S}_0: A, C_f, \top_f
\]

\[
C \subseteq B, \\
\mathcal{S}_0: C, B
\]

\[
s_0: A_e \cap \lnot B_f, \\
A_e, \lnot B_f
\]

\[
\mathcal{I}, \sqcup\text{-rules} \\
\mathcal{K} \models d: A_e \subseteq B_f
\]
Deduction example VIII

\[ A \equiv C_f, \]
\[ s_0 : A, C_f, T_f, B_f \]

\[ C \sqsubseteq B, \]
\[ s_0 : C, B \]

\[ s_0 : A_e \sqcap \neg B_f, \]
\[ A_e, \neg B_f, B_f \]

**A-rule**

\[ \mathcal{K} \models d : A_e \sqsubseteq B_f \]
Correctness and complexity

Correctness
Given CKR $\mathcal{R}$, $d \in D_\Gamma$, and concept $C$ on input,
- $C_T$ always terminates
- $C$ is $d$-satisfiable w.r.t. $\mathcal{R} \iff C_T$ generates complete and clash-free c.tree.

Idea: construction of CKR model from / to completion tree

Complexity
Complexity of $C_T$ is \text{NEXPTime} w.r.t. combined size of input.

Idea: by blocking, number of nodes limited by size of input $\mathcal{R}$ and $C$
1. Introduction

2. Contextualized Knowledge Repositories

3. Tableaux Algorithm

4. Algorithm Optimization

5. Conclusions
Optimization 1: precomputation of metaknowledge

- Rules use metaknowledge by querying for $d \prec e$
- Constant number of dimensions $k$ and contexts $m$
Optimization 1: precomputation of metaknowledge

- Rules use metaknowledge by querying for $d \prec e$
- Constant number of dimensions $k$ and contexts $m$

**Opt.1: precompute context structure**

- Idea: $k \times m^2$ queries to compute $\mathcal{M} \models d \prec_A d'$
- Meta-level reasoning does not slow down object-level reasoning
Optimization 2: parallelization (sketch)

- Contexts divide KB in smaller units (mostly independent)
- Part of reasoning can be parallelized (one processor per context)
Optimization 2: parallelization (sketch)

- Contexts divide KB in smaller units (mostly independent)
- Part of reasoning can be parallelized (one processor per context)

**Opt.2: parallelization (sketch)**

- **Local ($\exists$, $\exists$...)** & **top rules**: apply locally by each processor
- **$\Delta\uparrow, \Delta\downarrow$**: detect locally, send message to target context, cache individuals
- **$A, R$-rule**: detect by cache, send message to target context
- **$M$-rule**: detect locally, send message to target context
Optimization 3: propagation limitation

- Propagation: augments number of messages and computation
- Should be limited (also for parallelization)
Optimization 3: propagation limitation

- Propagation: augments number of messages and computation
- Should be limited (also for parallelization)

Opt.3: propagation limitation

- Lazy unfolding for $T_A, T_R, T_\subseteq \rightarrow T^*_A, T^*_R, T^*_\subseteq$
- Avoids introduction of $\neg T_d \sqcup T_e$ by $T_\subseteq$

Note: $C_T$ extended with $T^*_A, T^*_R, T^*_\subseteq$ is correct
Optimization 3: propagation limitation

\[ \mathcal{L}_d(x) = \{A_e\}, \ e \preceq f \]

\[ \mathcal{T}_A \]

\[ \mathcal{L}_d(x) = \{A_e, \top_f\}, \ e \preceq f \]

\[ \mathcal{L}_d(x) = \{\ldots\} \]
\[ \mathcal{L}_d(<x,y>) = \{R_e\} \]
\[ \mathcal{L}_d(y) = \{\ldots\} \]
\[ e \preceq f \]

\[ \mathcal{T}_R \]

\[ \mathcal{L}_d(x) = \{\top_f\} \]
\[ \mathcal{L}_d(<x,y>) = \{R_f\} \]
\[ \mathcal{L}_d(y) = \{\top_f\} \]
\[ e \preceq f \]

\[ \mathcal{L}_d(x) = \{\neg\top_f\}, \ e \prec f \]

\[ \mathcal{T}_\sqsubseteq \]

\[ \mathcal{L}_d(x) = \{\neg\top_f, \neg\top_e\}, \ e \prec f \]
Conclusions

**CKR:** contextual layer for DL knowledge bases

**$C_T$ tableaux algorithm:**
extension of $\text{NExpTime ALC}$ algorithm with CKR rules

**Optimizations:**
- metaknowledge precomputation
- parallelization (sketch)
- propagation limitation

**Reasoning: current and future directions**

- $\text{ExpTime}$ algorithm for $\text{ALC}$ CKR (submitted)
- Parallel version of $C_T$
- Optimization of $A, R$-rules
- Extend to $\text{SROIQ}$-based CKR
Thank you for listening

Towards More Effective Tableaux Reasoning for CKR

Loris Bozzato, Martin Homola, Luciano Serafini

https://dkm.fbk.eu/

DKM - Data and Knowledge Management Research Unit,
FBK-Irst, Fondazione Bruno Kessler - Trento, Italy


