INTEGRATING LOGICAL AND SUB-SYMBOLIC CONTEXTS OF REASONING *

Tarek Besold and Stefan Mandl
Chair of Computer Science 8: Artificial Intelligence, University of Erlangen-Nuremberg
Haberstraße 2, D-91058 Erlangen, Germany

Keywords: Information integration, Knowledge-based systems, Nonmonotonic reasoning, Knowledge representation.

Abstract: We propose an extension of the heterogeneous multi-context reasoning framework by G. Brewka and T. Eiter, which, in addition to logical contexts of reasoning, also incorporates sub-symbolic contexts of reasoning. The main findings of the paper are a simple extension of the concept of bridge rules to the sub-symbolic case and the concept of bridge rule models that allows for a straightforward enumeration of all equilibria of multi-context systems. We illustrate our approach with two examples from the fields of text and image classification.

1 INTRODUCTION

One of the important problems in knowledge representation and knowledge engineering is the impossibility of writing globally true statements about realistic problem domains. A circumstance that is also documented by the use of contexts and micro-theories in CYC ((Lenat, 1995)). Multi-context systems (MCS) are a formalization of simultaneous reasoning in multiple contexts . Different contexts are inter-linked by bridge rules which allow for a partial mapping between formulas/concepts/information in different contexts. Recently there have been a number of investigations of MCS reasoning (for instance, see (Roelofsen and Serafini, 2005) or (Brewka et al., 2007)), with (Brewka and Eiter, 2007) being one of the latest contributions. There, the authors describe reasoning in multiple contexts that may use different logics locally. Logical reasoning on the one hand is a special case of symbolic reasoning where, according to (Kurfess, 2002), entities of the application domain are represented by symbols. In sub-symbolic reasoning on the other hand domain entities are represented by (micro-)features.

There is no strict boundary between symbolic and sub-symbolic: what in one example are micro-features can be declared as entities and be symbolically reasoned about in another example (and vice versa). In this paper, we integrate contexts of logical reasoning and contexts of sub-symbolic reasoning into a single MCS. Possible applications of such reasoners are numerous. I.e. shortcomings of statistical methods could be remedied with declarative knowledge and vice versa.

2 INTEGRATING LOGICAL AND SUB-SYMBOLIC CONTEXTS OF REASONING

We now generalize concepts from (Brewka and Eiter, 2007) to be applicable for both logical and sub-symbolic reasoners.

Observation 1. *The concept of 'logic' as defined in (Brewka and Eiter, 2007) is – besides its name – also a valid characterization of sub-symbolic reasoning: The knowledge base consists of the available evidence (the input of the sub-symbolic reasoner), the set of possible belief sets is the set of possible results of the sub-symbolic reasoner, and the function* **ACC** *defines the actual reasoning. (It is* **ACC** *that is often generated from training examples.)*

In fact the only conceptual generalization that we make is to make no assumptions about the form of

*This is a short version of the paper. The full version, containing a short introduction to the notion of MCS and another more sophisticated application example from the domain of image classification, as well as proofs for the propositions stated, a categorization of our work and a comparison to similar approaches in the field can be found under `http://www8.informatik.uni-erlangen.de/inf8/Publications/bridging_mcs_original.pdf`.

Besold T. and Mandl S. (2010).
INTEGRATING LOGICAL AND SUB-SYMBOLIC CONTEXTS OF REASONING.
In *Proceedings of the 2nd International Conference on Agents and Artificial Intelligence - Artificial Intelligence*, pages 494-497
DOI: 10.5220/0002704704940497

the inputsof the reasoner which in the original work were assumed to be sets. Instead, we require test and update functions. This idea leads to our definition of a reasoner.

Definition 1. *A* reasoner *is a 5-tuple* $R = (\mathbf{Inp}_R, \mathbf{Res}_R, \mathbf{ACC}_R, \mathbf{Cond}_R, \mathbf{Upd}_R)$ *where* $\mathbf{Inp}_R$ *is the set of possible* inputs *to the reasoner,* $\mathbf{Res}_R$ *is the set of possible* results *of the reasoner,* $\mathbf{ACC}_R : \mathbf{Inp}_R \mapsto 2^{\mathbf{Res}_R}$ *defines the actual reasoning (assigning each input a set of results in a decidable manner),* $\mathbf{Cond}_R$ *is a set of decidable conditions on inputs and results,* $cond_R : \mathbf{Inp}_R \times \mathbf{Res}_R \mapsto \{0,1\}$*, and* $\mathbf{Upd}_R$ *is a set of update functions for inputs,* $upd_R : \mathbf{Inp}_R \mapsto \mathbf{Inp}_R$*.*

The example below shows that our Definition 1 comprises logics in the sense of (Brewka and Eiter, 2007) and sub-symbolic reasoners like Neural Nets, which originally were not covered. Hence Definition 1 is a generalization.

Example 1. *A logic L defined over a signature* Σ *(as of Definition 1) is a reasoner with* $\mathbf{Inp}_R = \mathbf{KB}_L$, $\mathbf{Res}_R = \mathbf{BS}_L$, $\mathbf{ACC}_R = \mathbf{ACC}_L$, $\mathbf{Upd}_R = \{\mathrm{fn}_x : kb \to kb \cup \{x\} | \forall x \in \bigcup_{k \in KB_L} k, \forall kb \in KB_L\}$, $\mathbf{Cond}_R = \{\mathrm{fn}_x : (\cdot, b) \to 1 \text{ iff } x \in b, 0 \text{ else} | x \in \bigcup_{bs \in BS_L} bs\}$

Example 2. *A standard feed-forward neural network* $\mathcal{N}$ *with n real valued inputs and m real valued outputs is a reasoner with* $\mathbf{Inp}_R = \mathbb{R}^n$, $\mathbf{Res}_R = \mathbb{R}^m$ *and* $\mathbf{ACC}_R(inp) = \{\mathcal{N}(inp)\}$ *where* $inp \in \mathbf{Inp}_R$*. In this case,* $\mathbf{ACC}_R$ *maps to singleton sets.* $\mathbf{Cond}_R$ *is a set of indicator functions on feature vectors,* $\mathbf{Upd}_R$ *is a set of update functions, each performing an update for a certain value of a component of a feature vector.*

The following definitions adapt the basic concepts of multi-context reasoning given in (Brewka and Eiter, 2007) for the use with reasoners as of Definition 1. In order to adapt the concept of bridge rules, we have to take into account the fact that the assumption of the reasoner inputs being sets is not made for general reasoners. Instead we have to use the defined test and update functions.

Definition 2. *Let* $R = \{R_1, \ldots, R_n\}$ *be a set of reasoners. An* R_k *-bridge rule over* $R, 1 \le k \le n$*, containing m conditions, is of the form*

$$u \leftarrow (r_1 : c_1), \ldots, (r_j : c_j), \mathbf{not}(r_{j+1} : c_{j+1}), \ldots, \mathbf{not}(r_m : c_m) \tag{1}$$

where $j \le m$, $1 \le r_k \le n$ *and* c_k *is a condition of inputs and results of some* R_{r_k} *and u is an element of* $\mathbf{Upd}_k$*.*

Definition 3. *A generalized multi-context system* $M = (C_1, \ldots, C_n)$ *consists of a collection of contexts* $C_i = (R_i, inp_i, br_i)$*, where* $R_i = (\mathbf{Inp}_i, \mathbf{Res}_i, \mathbf{ACC}_i, \mathbf{Cond}_i, \mathbf{Upd}_i)$ *is a reasoner,* inp_i *an input (an element of* $\mathbf{Inp}_i$ *), and* br_i *is a set of* R_i*-bridge rules over* $\{R_1, \ldots, R_n\}$ *as of equation (1).*

Concerning the belief states, we require input-output pairs instead of belief sets in every context.

Definition 4. *Let* $M = (C_1, \ldots, C_n)$ *be a generalized MCS. A generalized belief state is a sequence* $S = (S_1, \ldots, S_n)$ *such that each* S_i *is of the form* (inp_i, res_i) *with* $inp_i \in \mathbf{Inp}_i$ *and* $res_i \in \mathbf{Res}_i$*.*

We say a bridge rule r of form (1) is applicable in a generalized belief state $S = (S_1, \ldots, S_n)$ iff for $1 \le i \le j : c_i(inp_i, res_i) = 1$ in S_i and for $j+1 \le k \le m : c_k(inp_k, res_k) = 0$ in S_k. Now we prepare for the concept of equilibrium in the generalized setting.

Definition 5. *The set of (context local) update functions with respect to a corresponding element* S_i *of a belief state S is given by* $\mathbf{US}_i(MCS, S) = \{head(r) | r \in br_i \text{ applicable in } S\}$*, where* br_i *denotes the set of bridge rules of* S_i*'s corresponding context* C_i*.*

In general, a set of update functions may yield different results when the functions are applied multiple times or in different orders. We do not allow such sets of update functions.

Definition 6. *An applicable set of update functions* $\mathbf{US}_i(MCS, S)$ *is stationary for an input* inp_i *iff the following two conditions hold:* $\forall u \in \mathbf{US}_i(MCS, S) : u(inp_i) = u^m(inp_i)$ *for* $m \ge 1$ *(i. e. idempotency), and* $\forall u, u' \in \mathbf{US}_i(MCS, S) : u(u'(inp_i)) = u'(u(inp_i))$ *(i. e. commutativity).*

Definition 7. *The update of a belief state element* S_i *of a belief state S, with respect to a set of update functions* $\mathbf{US}_i$ *with k elements, is given by* $u_1(u_2(\ldots u_k(inp_i)\ldots))$ *if* $\mathbf{US}_i$ *is stationary for* inp_i*, and undefined otherwise.*

Please note that stationarity is only required for the set of update functions that is actually applied to belief state elements at a time. We now can give the definition of the generalized concept of equilibrium.

Definition 8. *A generalized belief state* $S = ((inp_1, res_1), \ldots, (inp_n, res_n))$ *of M is an equilibrium iff, for* $1 \le i \le n$*, the following condition holds:* $update(inp_i, \mathbf{US}_i) = inp_i$ *and* $res_i \in ACC(inp_i)$ *where* $update(inp_i, \mathbf{US}_i)$ *denotes the update of* S_i *with respect to* US_i*, which in turn has to be stationary for the corresponding* inp_i*.*

Please note that inputs for which the update is non-stationary are not part of any equilibrium.

Proposition 1. *Definitions 1 to 8 are a generalization of Definitions 1 to 5 in (Brewka and Eiter, 2007).*

3 COMPUTING EQUILIBRIA FOR FINITE MCS

For a belief state being an equilibrium only means that all the bridge rules are respected. As local reasoners may be non-monotonic and, furthermore, the bridge rules are non-monotonic also, there may be several equilibria for a given MCS. In general it is not clear which one constitutes the desired one.

When no external knowledge about preferences (e. g. a preference function which induces an order on equilibria) is available, in the field of computational logic, there exists the *principle of minimality*. Sadly minimality in general has no straightforward translation to sub-symbolic reasoning contexts (e. g. for vector valued sub-symbolic reasoning contexts we would need some kind of metric).

As for a deterministic sub-symbolic reasoner, given a set of inputs, there is exactly one corresponding set of results, one may nonetheless try to carry over minimality from the symbolic-only background. Using the notion of C*-minimality as introduced by in (Brewka and Eiter, 2007), minimality may be demanded for the symbolic contexts of a generalized MCS, which may in this case be composed of symbolic and deterministic sub-symbolic reasoners. As the deterministic sub-symbolic reasoners only yield exactly one set of results for a given set of inputs (and no phenomena as self-sustaining equilibria are possible), the C*-minimality generalizes to a global property of the equilibrium.

The remainder of this section describes a procedure to compute all equilibria of a finite MCS, based on complete enumeration. Thus criteria as e. g. minimality may be applied to the set of equilibria afterwards. Part of future research will be to construct more specialized algorithms, already exploiting the properties of ordering relations during the computation.

Definition 9. *An MCS* $M = (C_1, \ldots, C_n)$ *is said to be finite, iff for* $1 \leq i \leq n$*, following condition holds:* $|ACC(inp_i)| < \infty$ *and* $|br_i| < \infty$*.*

For the implementation, we consider finite MCS only.

Definition 10. *Let* **Br** *be a set of n bridge rules of an MCS. A bridge rule model is an assignment* $\mathbf{Br} \mapsto \{0,1\}^n$ *that represents for each bridge rule in* **Br** *whether it is active or not.*

Proposition 2. *For each equilibrium there is exactly one bridge rule model.*

For a given bridge rule model and an MCS we first apply all the bridge rules activated in the bridge rule model yielding $inp'_1 ... inp'_n$. Then we compute the set of results for each context i given inp'_i by applying $ACC(inp'_i)$, yielding a set of results res_i^j for each i, being of finite cardinality as MCS was said to be finite. Thus, testing whether (inp_i, res_i^j) is an equilibrium for all j, we obtain the set of equilibria for the given bridge rule model. Iterating the procedure over the (finite) set of all bridge rule models and joining the resulting sets of equilibria finally yields the set of all equilibria.

Definition 11. *Given an MCS with a (global) set of bridge rules* $br = \bigcup_i br_i$*. A set of bridge rules* $br_j \subseteq br$ *is called update-monotonic iff for all belief states S, S' the following condition holds:* $S' = update(MCS, S) \Rightarrow VC(MCS, S) \subseteq VC(MCS, S')$ *where* $VC(MCS, S) = \bigcup_i \{cond_i \in R_i | cond_i(inp_i, res_i) = 1\}$ *and* $update(MCS, S)$ *is the (global) update over all* $S_i \in S$*.*

As bridge rules in the update-monotonic subset of bridge rules of the MCS are guaranteed to remain active after any update, the update-monotonic bridge rules that are initially active in the MCS when searching for equilibria have to be active in any equilibrium. Hence, when iterating over all bridge rule models, only those bridge rule models that comply with the initially active update-monotonic bridge rules have to be considered.

As a downside computing the update-monotonic subset of the bridge rules depends on the idiosyncrasies of the reasoners involved, condition test and update functions and therefore cannot be performed in general. Another inconvenience is the fact that if there are no update-monotonic bridge rules all elements of the entire set of bridge rule models have to be tested for representing an equilibrium.

Definition 12. *A reasoner* $R = (\mathbf{Inp}_R, \mathbf{Res}_R, \mathbf{ACC}_R, \mathbf{Cond}_R, \mathbf{Upd}_R)$ *is deterministic iff* $\mathbf{ACC}_R(x)$ *is a singleton set for every* $x \in \mathbf{Inp}_R$*.*

Proposition 3. *For an MCS with deterministic reasoners only, there exists at the most one equilibrium for each bridge rule model.*

Applying the proposition to the algorithm sketched above, one may reduce the number of pairs (inp_i, res_i^j) to be tested for being an equilibrium, by testing each pair (inp_i, res_i^j) directly after it was generated, and switching to the next bridge rule model after having found an equilibrium for this very model, as there may be one at the most.

4 EXAMPLE

We present an example application for multi-context reasoning which involves logical and sub-symbolic contexts of reasoning. As logical reasoners, we assume a propositional logical reasoner R_{pl}. As sub-symbolic reasoners, we assume a maximum-likelihood reasoner using the Naive Bayes assumption R_{nb}.

4.1 Text Classification

This example is taken from the domain of (statistical) text classification. Text may be categorized into two different classes: 'music event' and 'political event'. We assume a binomial text model (see (Manning et al., 2008) for more details) and – for the sake of simplicity – a vocabulary of only two terms: *Queen, Elizabeth*.

Term t	$\mathbf{P}(\text{mus. ev.}\|t)$	$\mathbf{P}(\text{pol. ev.}\|t)$
Queen	4.1×10^{-3}	2.0×10^{-3}
Elizabeth	3.0×10^{-3}	1.9×10^{-2}
Query q	Naive Bayes probability of q	
Queen Elizabeth	1.2×10^{-5}	3.8×10^{-5}

Figure 1: Prior probabilities of terms and Naive Bayes probability of a query for the classes 'music event' and 'political event'.

Figure 1 shows the prior probabilities of the classes given the terms, and the Naive Bayes probabilities given the combined query *Queen Elizabeth*. The priors have been obtained by querying a web search engine, but for the example the actual source of the priors is not of much interest. Those probabilities define the function $\mathbf{ACC}_{nb}$ of R_{nb}. When stating the query '*Queen*' to the reasoner ($inp_{nb} = \{Queen\}$), the result (via maximum likelihood) is $res_{nb} = \{$'music event'$\}$.

We would like to improve the reasoner R_{nb} by providing specific knowledge about the British Royals. Hence, we use a proposition logical reasoner R_{pl} together with a knowledge base $inp_{pl}=\boxed{Elizabeth \leftarrow Queen}$ of relevant information.

In order to link the two reasoners, we define bridge rules $br_{pl} = \{add\_Queen \leftarrow nb : has\_input\_Queen\}$ and $br_{nb} = \{add\_to\_query\_Elizabeth \leftarrow pl : holds\_Elizabeth\}$, where the condition and update functions have the obvious meaning. Taking the query '*Queen*' into account, the MCS for reasoning in R_{nb} and R_{pl} is given by $M^t = \{C_{nb}, C_{pl}\}$ with $C_{nb} = (R_{nb}, \{Queen\}, br_{nb})$, $C_{pl} = (R_{pl}, inp_{pl}, br_{pl})$.

Then, the belief state

$$((\{Queen, Elizabeth\}, \{\text{'political event'}\}), (inp_{pl} \cup \{Queen\}, inp_{pl} \cup \{Queen, Elizabeth\}))$$

is the only equilibrium of M^t. Hence, with multi context reasoning, the result for the query '*Queen*' has been changed from 'music event' to 'political event'.

5 CONCLUSIONS

The paper presents a generalization of heterogeneous multi-context systems that allows for the use of sub-symbolic contexts of reasoning alongside logical contexts of reasoning. An exhaustive algorithm for enumerating all equilibria of an MCS is given.

Still, the lack of a conceptual notion of minimality or stability for sub-symbolic beliefs poses a challenge for future research, which we are confident to handle in the near future.

On the pragmatic side, the illustrative examples demonstrate that a more powerful language to describe updates and conditions on reasoner inputs and results, respectively, has to be developed in order to allow for concise definitions of bridge rules.

REFERENCES

Brewka, G. and Eiter, T. (2007). Equilibria in heterogeneous nonmonotonic multi-context systems. In *Proceedings of the National Conference on Artificial Intelligence*, volume 22.

Brewka, G., Roelofsen, F., and Serafini, L. (2007). Contextual Default Reasoning. *Proc. IJCAI-07, Hyderabad, India*.

Kurfess, F. J. (2002). Integrating symbol-oriented and sub-symbolic reasoning methods into hybrid systems. In *Proceedings of the International School on Neural Networks. Fifth Course: From Synapses to Rules: Discovering Symbolic Rules from Neural Processed Data*.

Lenat, D. B. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38.

Manning, C. D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Roelofsen, F. and Serafini, L. (2005). Minimal and Absent Information in Contexts. In *International Joint Conference on Artificial Intelligence*, volume 19, page 558. Lawrence Erlbaum Associates LTD.