Graph-Embedded Term Rewrite Systems and Applications (A Preliminary Report)

Saraid Dwyer Satterfield[1], Serdar Erbatur[2], Andrew M. Marshall[1], and Christophe Ringeissen[3]

[1] University of Mary Washington, USA
[2] University of Texas at Dallas, USA
[3] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France

1 Introduction

In this preliminary paper we report on the development of a new form of term rewrite system, called the graph-embedded term rewrite systems, and motivate the study and use of such rewrite systems by demonstrating their usefulness in the application of security protocols.

The research area of cryptographic protocol analysis contains a number of innovative algorithms and procedures for checking various security properties of protocols, see for example [1, 3, 7, 9, 11]. These procedures consider protocols modeled in a symbolic way, typically via a rewrite system or equational theory. Often the procedure is proven sound and complete for specific classes of theories. One of the most common classes are those theories that can be represented by subterm convergent term rewrite systems. That is, term rewrite systems where the right-hand side of the rules are constants or strict subterms of the left-hand side. For example, see the procedures developed in [1, 9]. Interestingly, many of these same procedures also work for theories that are "beyond subterm", that is they are not strictly subterm convergent. However, since these examples don't fit into a known class of theories for which soundness and completeness proofs already exist, they must be proven on an individual bases. For example, the procedures of [1, 9] are shown to work on the theory of blind signatures, see Example 1 below. However, the theory is not subterm convergent, notice in the final rule, $unblind(sign(blind(x, y), z), y) \rightarrow sign(x, z)$, that $sign(x, z)$ is not a strict subterm of $unblind(sign(blind(x, y), z), y)$. Thus, in each case a unique proof is needed to show applicability of the procedure on the theory of blind signatures. Several additional examples of beyond subterm theories are given throughout the paper. This begs the question of whether there is a syntactic definition of a class of term rewrite systems such that the definition encapsulates these beyond subterm examples yet still maintains some of the useful properties needed to ensure applicability of the above procedures.

This paper has begun the exploration of the above question by introducing and studying the new notion of graph-embedded term rewrite systems. The graph embedded systems encompass most of the beyond subterm examples from many of the protocol analysis procedures [1,7,9,11]. As an initial step, in this paper we concentrate on the procedure developed in [1] and the notion of locally stable theories, an important property in the decidability of deducibility. With the new definition of graph-embedded system we are now able to easily identify a symbolic class of term rewrite system which are beyond subterm, encompass most of the beyond subterm examples of [1, 7, 9, 11], and are locally stable.

Finally, this paper represents the initial exploration of graph-embedded term rewrite systems and their application to protocol analysis. We hope that the formulation proves useful in areas beyond security protocols as homeomorphic embeddings have proven useful in many areas.

2 Preliminaries

We use the standard notation of equational unification [5] and term rewriting systems [4]. The size of a term t is denoted by $|t|$ and defined in the usual way as follows: $|f(t_1, \ldots, t_n)| = 1 + \Sigma_{i=1}^{n} |t_i|$ if f is a n-ary function symbol with $n \geq 1$, $|c| = 1$ if c is a constant, and $|x| = 1$ if x is a variable. Let $VP(t)$ denote the set of leaf nodes in the term-graph of a term t labeled by a variable. Notice that two distinct positions could be label by the same variable. Let $FP(t)$ denote the set of nodes in the term-graph of t labeled by a function symbol. Notice that two distinct positions could be labeled by the same function symbol. Let $FS(t)$ denote the set of function symbols in the term t. We write $t \approx t'$ if the term t is equal to the term t' modulo a permutation of the leaf nodes. Recall that we can define the Homeomorphic embedding relation, $\trianglerighteq_{emb}$, as the reduction relation $\rightarrow^*_{R_{emb}}$ induced by the following rewrite system: $R_{emb} = \{f(x_1, \ldots, x_n) \rightarrow x_i \mid n \geq 1,\ 1 \leq i \leq n\}$

Example 1. *The theory of blind signatures [9], is an example of a TRS for which each right-hand side is a homeomorphic embedding of the left-hand side:* $open(commit(x, y), y) \rightarrow x$, $getpk(host(x)) \rightarrow x$, $checksign(sign(x, y), pk(y)) \rightarrow x$, $unblind(blind(x, y), y) \rightarrow x$, $unblind(sign(blind(x, y), z), y) \rightarrow sign(x, z)$.

Notions of Knowledge. The applied pi-calculus and frames are used to model attacker knowledge [2]. In this model, the set of messages or terms which the attacker knows are the set of terms in $Ran(\sigma)$ of the frame $\phi = \nu\tilde{n}.\sigma$, where σ is a substitution ranging over ground terms. $\tilde{n}$ consists of a finite set of restricted names which remain secret from the attacker. The set of names occurring in a term t is denoted by $fn(t)$. For any frame, $\phi = \nu\tilde{n}.\sigma$, $Dom(\phi) = Dom(\sigma)$. We recall the characterization of deduction from [1] that if M is a closed term and $\phi = \nu\tilde{n}.\sigma$ a frame. Then, $\phi \vdash M$ iff there exists a term ζ s.t. $fn(\zeta) \cap \tilde{n} = \emptyset$ and $\zeta\sigma =_E M$.

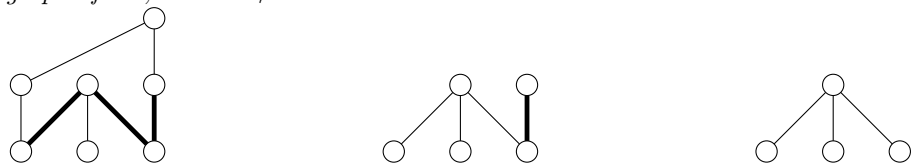
Some Graph Theory If $G = (V, E)$ and $G' = (V', E')$ are two graphs, then G and G' are isomorphic, $G \simeq G'$, if there exists a bijection $\phi : V \rightarrow V'$ with $xy \in E$ iff $\phi(x)\phi(y) \in E'$, $\forall x, y \in V$.

Definition 1. *Let $G = (V, E)$ and $e = xy$. G/e is the graph $G' = (V', E')$ such that $V' = (V \setminus \{x, y\}) \cup \{v_e\}$, where v_e is a new vertex, and $E' = \{vw \in E \mid \{v, w\} \cap \{x, y\} = \emptyset\} \cup \{v_e w \mid xw \in E \setminus \{e\}$ or $yw \in E \setminus \{e\}\}$. We say that G' is obtained from G by* contracting *the edge e.*

Definition 2. *[10] G is an MG', denoted $G = MG'$, if G' can be obtained from G by a series of edge contractions. That is, iff there exists graphs $G_0, G_1, \ldots, G_n$ and edges $e_i \in G_i$ such that $G = G_0$, $G_n \simeq G'$, and $G_{i+1} = G_i/e_i$ for all $i < n$. If $G = MG'$ and G is a subgraph of another graph G'', we call G' a* graph minor *of G'', denoted as $G' \succcurlyeq G''$.*

The above type of embedding, G' in G'', provides flexibility while preserving required features. In particular, when translated to terms, it will preserve a more flexible type of subterm.

Example 2. *Notice that G' is obtained from G via a edge contraction. In addition G is a subgraph of G'', thus $G' \succcurlyeq G''$.*



(G'') (G) (G')

3 Graph-Embedded Systems

The key to translating from the graph theory setting to the term setting is to use the same methods, contractions, but require that the *final* term graph constructed in this fashion represent a *well-formed* term. That is, we need to enforce the notion of a well formed term. We develop a set of rewrite *schema* which preserve a type of graph minor relation. This set of schema then induces a graph-embedded term rewrite system. This is very similar to what is done with homeomorphic embeddings.

Definition 3. *Consider the following reduction relation, $\rightarrow^*_{R_{gemb}}$, induced by the set of rewrite rules create after instantiating the following rule schema with Σ:*

$$R_{gemb} = \left\{ \begin{array}{ll} (1) & f_j(x_1, \ldots, x_n) \rightarrow x_i \\ (2) & f_j(x_1, \ldots,\ x_{i-1},\ x_i,\ x_{i+1}\ \ldots,\ x_n) \rightarrow f_j(x_1, \ldots,\ x_{i-1},\ x_{i+1},\ \ldots,\ x_n) \\ & \text{and for any } f_j, f_k \in \Sigma \\ (3) & f_j(x_1, \ldots, x_{i-1}, f_k(\bar{z}), x_{i+1},\ \ldots, x_m) \rightarrow f_k(x_1, \ldots,\ x_{i-1}, \bar{z}, x_{i+1},\ \ldots, x_m) \\ (4) & f_j(x_1, \ldots, x_{i-1}, f_k(\bar{z}), x_{i+1},\ \ldots, x_m) \rightarrow f_j(x_1, \ldots,\ x_{i-1}, \bar{z}, x_{i+1},\ \ldots, x_m)\} \end{array} \right\}$$

We say a term t' is graph embedded in a term t, denoted $t' \succcurlyeq_{gemb} t$, if t' is a well formed term and $t \rightarrow^*_{R_{gemb}} s \approx t'$.

Remark 1. *Notice that the rules in R_{gemb} ignore function arity, thus intermediate terms between t and t' may not be well formed. It is only the final term for which function arity and the relation between variables and functions must obey the standard term definition requirements.*

We can now introduce the notion of graph-embedded term rewrite system.

Definition 4. *A TRS R is a graph embedded TRS if $\forall\ l \rightarrow r \in R$, $r \succcurlyeq_{gemb} l$.*

Example 3. *Theory of malleable encryption is defined by $R_{mal} = \{dec(enc(x,y),y) \rightarrow x$, $mal(enc(x,y),z) \rightarrow enc(z,y)\}$. For the final rule, $mal(enc(x,y),z) \rightarrow_{R_{gemb}} enc(x,y,z) \rightarrow_{R_{gemb}} enc(y,z) \approx enc(z,y)$. Thus $\forall l \rightarrow r \in R_{mal}$, $r \succcurlyeq_{gemb} l$. Notice that $enc(x,y,z)$ is not a well formed term since it violates the arity of $enc()$. However, the final term is well formed, as required.*

Example 4. *Theory of trap-door commitment, from [9], is also graph-embedded:*
$R_{tdc} = \{open(td(x,y,z),y) \rightarrow x$, $open(td(x,y,z), f(x_1,y,z,x_2)) \rightarrow x_2$,
$td(x_2, f(x_1,y,z,x_2),z) \rightarrow td(x_1,y,z)$, $f(x_2, f(x_1,y,z,x_2),z,x_3) \rightarrow f(x_1,y,z,x_3)\}$.

Example 5. *The theory of blind signatures from Example 1 is also a graph-embedded TRS. All but the final rule are subterm. For the final rule, $unblind(sign(blind(x,y),z),y) \rightarrow_{R_{gemb}} sign(blind(x,y),z)$ via rule (1). $sign(blind(x,y),z) \rightarrow_{R_{gemb}} sign(x,y,z)$ via rule (3). Notice again that this intermediate term is not well formed. Finally $sign(x,y,z) \rightarrow_{R_{gemb}} sign(x,z) \approx sign(x,z)$ via rule (2).*

Similar to the class of subterm TRSs, the graph-embedded TRSs have several nice properties such as termination. They are also useful for modeling many beyond subterm theories found in security protocols, we explore that application next.

4 Applications to Security Protocols

In this section we focus on theories with the local stability property as introduced in [1]. For this purpose, it's useful to consider a restricted form of graph-embedded.

Definition 5. *Let R be a graph embedded TRS such that for all $l \to r \in R$ one of the following is true: r is a variable or constant, or: $l \to^*_{R_{gemb}} r' \approx r$ s.t. R_{gemb} consists of all the rules of Definition 3 except rule (3) is replaced by the following rule; $f_j(x_1, \ldots, x_{i-1}, f_k(\bar{z}), x_{i+1}, \ldots, x_m) \to f_k(\bar{z})$. Then, R is a* contracting graph-embedded *TRS.*

The motivation behind this definition is to disallow moving layers of a term that are at the same or higher positions as a function symbol under that symbol.

Example 6. *Consider the following TRSs. The theory of blind signatures from Example 1 is a contracting graph-embedded TRS. The theory of trap-door commitment from Example 4 is a contracting graph-embedded TRS. The theory of malleable encryption from Example 3 is not a contracting graph-embedded TRS. Notice that for the rule $mal(enc(x, y), z) \to enc(z, y)$, the node labeled with z is moved under the enc node on the right-hand side. This violates the requirements of Definition 5. The theory, consisting of the rule $f(g(x)) \to g(h(x))$, from [9], which is also non-terminating in the procedure from [9], is not graph-embedded and thus not contracting graph-embedded.*

Remark 2. *One could naturally ask if the above definition just leads to systems with the Finite Variant Property (FVP), another useful property for some decision procedures [7, 11]. However, one can easily construct contracting graph-embedded systems that do not have the boundedness property (a TRS has the FVP iff it has the boundedness property [6, 8, 12]). For example, $R = \{f(h(x)) \to f(x)\}$.*

We need to introduce a few notions needed when considering local stability and computing a saturation set for a frame. Let $ar(\Sigma)$ denote the maximal arity of any function symbol in Σ. Define the context bound of a graph-embedded TRS, $R = \{l_i \to r_i\}$, $1 \leq i \leq n$, as $c_R = max_{1 \leq i \leq n}(|l_i|, ar(\Sigma) + 1)$.

Example 7. *For the theory of malleable encryption from Example 3, $c_{R_{mal}} = 5$. For the theory of blind signatures, R_{blind}, from Example 1, $c_{R_{blind}} = 7$.*

Definition 6. *Let R be a contracting graph-embedded TRS and let $st(t)$ be the set of subterms of a term t. Then, the set of graph-embedded subterms of a term t, denoted as $gst(t)$, is defined as: $gst(c) = c$, where c is a name or a constant, $gst(t) = \{t' | t \to^*_{R_{gemb}} t'' \approx t'\} \cup \bigcup_{t'' \in st(t)} gst(t'')$. Let $\phi = \nu\tilde{n}.\{M_1/x_1, \ldots, M_k/x_k\}$ be a frame, then $gst(\phi) = \cup_{i=1}^{k}(gst(M_i))$.*

Notice that for any term t, $gst(t)$ is a finite set. This is due to the fact that when recursively constructing $gst(t)$ in the second rule of Definition 6, t' is strictly smaller than t, and any term $t'' \in st(t)$ must also be strictly smaller than t. Based on the extended definition of subterms, gst, we can now construct a saturation set for frames which is critical to modeling attacker knowledge and deciding deducibility [1].

Definition 7. *Let $\phi = \nu\tilde{n}.\{M_1/x_1, \ldots, M_k/x_k\}$ be a frame, and R a contracting graph-embedded TRS. Define the set $sat(\phi)$ to be the smallest set such that: (1) $M_1, \ldots, M_k \in sat(\phi)$, and $n \in sat(\phi)$ for every $n \in fn(\phi)$, (2) if $M_1, \ldots, M_l \in sat(\phi)$ and $f(M_1, \ldots, M_l) \in gst(\phi)$, then $f(M_1, \ldots, M_l) \in sat(\phi)$, (3) if $M_1, \ldots, M_l \in sat(\phi)$, $C[M_1, \ldots, M_l] \to^{\epsilon}_R M$, where C is a context, $|C| \leq c_R$, $fn(C) \cap \tilde{n} = \emptyset$, and $M \in gst(\phi)$, then $M \in sat(\phi)$.*

This set is also finite and this finiteness is critical to computing the possible attackers knowledge. In addition, it can be shown that if $M \in sat(\phi)$ for some frame ϕ, then $\phi \vdash M$. This is a required condition for the establishment of local stability. The following definition and two lemmas establish the remaining requirements for local stability.

Definition 8. *Let R be a TRS, C a context, and $S_1, S_2, \ldots, S_n \in gst(\phi)$, for some frame ϕ and $n > 0$. If whenever $C[S_1, \ldots, S_n] \rightarrow_R t$ we have that $t = C'[S'_1 \ldots S'_m]$ for some context C' and some $S'_1, \ldots, S'_m \in gst(\phi)$, then we say that $gst(\phi)$ is* closed under context by R.

Lemma 1. *Let $\phi = \nu\tilde{n}.\{M_1/x_1, \ldots, M_k/x_k\}$ be a frame. If R_{gemb} is the modified set of graph-embedded rules of Definition 5. Then, $gst(\phi)$ is closed under context by R_{gemb}.*

Lemma 2. *Let $\phi = \nu\tilde{n}.\{M_1/x_1, \ldots, M_k/x_k\}$ be a frame. Let R_{gemb} be the modified set of graph-embedded rules of Definition 5. Let $S_1, \ldots, S_l \in gst(\phi)$ and assume $C[S_1, \ldots, S_l] \rightarrow^*_{R_{gemb}} M' \approx M$, where C is some context and M is a well formed term. Then, $M = C'[S'_1, \ldots, S'_k]$, where $S'_1, \ldots, S'_k \in gst(\phi)$ and C' is a context with $|C'| \leq |C|$.*

One method for ensuring the decidability of deduction is for the saturation of a frame to produce a set that is closed under the application of "small" context. This condition, called locally stable, is introduced in [1] and improved in [3]. A simplified version of this definition is introduced below. It is simplified because we don't consider AC-symbols as in [1, 3].

Definition 9. *(locally stable [1]). A convergent TRS, R, is* locally stable *if, for every frame $\phi = \nu\tilde{n}.\{M_1/x_1, \ldots, M_k/x_k\}$, where each M_i is a closed term in R normal form, there exist a finite set $sat(\phi)$ such that: (1) $M_1, \ldots, M_k \in sat(\phi)$ and $n \in sat(\phi)$, for all $n \in fn(\phi)$; (2) if $M_1, \ldots, M_k \in sat(\phi)$ and $f(M_1, \ldots, M_k) \in st(sat(\phi))$, then $f(M_1, \ldots, M_k) \in sat(\phi)$; (3) if $C[S_1, \ldots, S_l] \rightarrow^{\epsilon}_R M$, where C is a context with $|C| \leq c_R$ and $fn(C) \cap \tilde{n} = \emptyset$, and each $S_i \in sat(\phi)$, then there exists a context C', and $S'_1 \ldots S'_k \in sat(\phi)$, such that $|C'| \leq c_R$, $fn(C') \cap \tilde{n} = \emptyset$, and $M \rightarrow^*_R M' = C'[S'_1, \ldots, S'_k]$; (4) if $M \in sat(\phi)$ then $\phi \vdash M$.*

We can now establish the main result.

Theorem 1. *Let R be a convergent contracting graph-embedded TRS. Then, R is locally stable.*

Combining the locally stable property with the property of *locally finite* allows for the decidability of not only deduction but static equivalence. Let $\phi = \nu\tilde{n}.\sigma$ be a frame and let $Eq(\phi)$ be the set of equalities, $s = t$, such that $s\sigma =_E t\sigma$ and $\tilde{n} \cap (fn(s) \cup fn(t)) = \emptyset$. The difficulty is that $Eq(\phi)$ is not always finite and computable. The property of *locally finite* says that $Eq(\phi)$ is always equivalent to another set that is finite and computable. See [1] for further discussion and the proof that convergent theories without AC symbols are locally finite. Static equivalence captures properties important to security protocols not captured by deducibility. Two frames, ϕ_1 and ϕ_2, are *static equivalent* when $Dom(\phi_1) = Dom(\phi_2)$ and $Eq(\phi_1) = Eq(\phi_2)$.

Example 8. *Continuing Example 6, since blind signatures is a contracting graph-embedded TRS, it is locally stable by Theorem 1. The theory of blind signatures doesn't contain an AC-symbol thus it is locally finite [1]. Therefore, blind signatures is both locally stable and locally finite and thus both deduction and static equivalence are decidable [1].*

Directly from Theorem 1 and the result in [1], we obtain the following corollary.

Corollary 1. *Let R be a convergent and contracting graph-embedded TRS. Then deduction and static equivalence are decidable for R.*

Finally, in continuing this line of research we would also like to consider termination conditions of various procedures [1, 7, 9, 11] for graph-embedded systems. We would also like to see if the graph-embedded idea could be extended. For example, not all theories considered in [1, 7, 9, 11] are graph embedding. It would be interesting to see if such systems could be considered or if additional graph theory concepts such as topological minors could be useful.

References

[1] Martín Abadi and Véronique Cortier. Deciding knowledge in security protocols under equational theories. *Theor. Comput. Sci.*, 367(1-2):2–32, 2006.

[2] Martín Abadi and Cédric Fournet. Mobile values, new names, and secure communication. In *Proceedings of the 28th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL'01, pages 104–115, New York, NY, USA, 2001. ACM.

[3] Mauricio Ayala-Rincón, Maribel Fernández, and Daniele Nantes-Sobrinho. Intruder deduction problem for locally stable theories with normal forms and inverses. *Theor. Comput. Sci.*, 672:64–100, 2017.

[4] Franz Baader and Tobias Nipkow. *Term rewriting and all that.* Cambridge University Press, New York, NY, USA, 1998.

[5] Franz Baader and Wayne Snyder. Unification theory. In John Alan Robinson and Andrei Voronkov, editors, *Handbook of Automated Reasoning*, pages 445–532. Elsevier and MIT Press, 2001.

[6] Christopher Bouchard, Kimberly A. Gero, Christopher Lynch, and Paliath Narendran. On forward closure and the finite variant property. In Pascal Fontaine, Christophe Ringeissen, and Renate A. Schmidt, editors, *Frontiers of Combining Systems*, volume 8152 of *Lecture Notes in Computer Science*, pages 327–342. Springer Berlin Heidelberg, 2013.

[7] Rohit Chadha, Vincent Cheval, Ştefan Ciobâcă, and Steve Kremer. Automated verification of equivalence properties of cryptographic protocols. *ACM Trans. Comput. Log.*, 17(4):23:1–23:32, 2016. Available as Research Report at https://hal.inria.fr.

[8] Hubert Comon-Lundh and Stéphanie Delaune. The finite variant property: How to get rid of some algebraic properties. In Jürgen Giesl, editor, *Rewriting Techniques and Applications*, volume 3467 of *Lecture Notes in Computer Science*, pages 294–307. Springer, 2005.

[9] Ştefan Ciobâcă, Stéphanie Delaune, and Steve Kremer. Computing knowledge in security protocols under convergent equational theories. *J. Autom. Reasoning*, 48(2):219–262, 2012.

[10] Reinhard Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, third edition.

[11] Jannik Dreier, Charles Duménil, Steve Kremer, and Ralf Sasse. Beyond subterm-convergent equational theories in automated verification of stateful protocols. In Matteo Maffei and Mark Ryan, editors, *Principles of Security and Trust*, pages 117–140, Berlin, Heidelberg, 2017. Springer Berlin Heidelberg.

[12] Santiago Escobar, Ralf Sasse, and José Meseguer. Folding variant narrowing and optimal variant termination. *J. Log. Algebr. Program.*, 81(7-8):898–928, 2012.