Carving Rule-based Filters within a Spatio-temporal Logic

Shyamanta M Hazarika
Computer Sc. & Engineering
Tezpur University
INDIA
smh@tezu.ernet.in

Acknowledgement

Work reported here is part of a **Collaborative Project** between TU, Tezpur and CDAC, Bangalore funded by DIT, Govt. of India.

Filters vis-à-vis Firewalls

- **Firewalls** are the frontier defense in network security, **filtering out unwanted packets** coming from or going to the secured network.

- Filtering is rule-based; in essence **firewall is a set of ordered filtering rules** configured primarily based on predefined security policy.

Filters vis-à-vis Firewalls

- Each **rule is formed of a condition** and **an action**. A rule condition is a set of fields; any field in IP, UDP or TCP headers may be used.
- Fields in a firewall rule could have a singular value or a range of values.
- Filtering **actions** are either to **accept or** to **deny**.

Intra-Firewall Anomalies

- Filtering policy within a firewall is dependent on the ordering of filtering rules.
  - Note that for a set of completely disjoint filter rules, the ordering is insignificant.
  - This is not usually the case and therefore ordering is important.
- Further, as number of rules increases, there are possibility of writing conflicting or redundant rules.

Intra-Firewall Anomalies

- Intra-firewall **policy anomaly** is the existence of
  - two or more filtering rules that may match the same packet or
  - a rule that can never match any packet that cross the firewall.
- **Anomalies** are properties of filters that hint at **possible misconfiguration** and have been well studied in the literature.
  - Traditional **anomaly-detection** algorithms have been shown to run in time **exponential** in the number of filter rules.

Filter Rules vis-à-vis Firewall

<order>	<protocol>	<s_ip>	<s_port>	<d_ip>	<d_port>	<action>
1:	tcp,	140.192.37.20,	any,	*.*.*.*,	80,	deny
2:	tcp,	140.192.37.*,	any,	*.*.*.*,	80,	accept
3:	tcp,	*.*.*.*,	any,	161.120.33.40,	80,	accept
4:	tcp,	140.192.37.*,	any,	161.120.33.40,	80,	deny
5:	tcp,	140.192.37.30,	any,	*.*.*.*,	21,	deny
6:	tcp,	140.192.37.*,	any,	*.*.*.*,	21,	accept
7:	tcp,	140.192.37.*,	any,	161.120.33.40,	21,	accept
8:	tcp,	*.*.*.*,	any,	*.*.*.*,	any,	deny
9:	udp,	140.192.37.*,	any,	161.120.33.40,	53,	accept
10:	udp,	*.*.*.*,	any,	161.120.33.40,	53,	accept
11:	udp,	140.192.38.*,	any,	161.120.35.*,	any,	accept
12:	udp,	*.*.*.*,	any,	*.*.*.*,	any,	deny

Motivation

- Filter **rules can be seen as 'spatial' regions**. Rules are applied sequentially. Such **sequential order of rules is 'temporal'**.

- Herein lies the appeal for **characterizing rule-based filters within a spatio-temporal logic**; we present an approach to carving rule-based filters within ST_0.

Qualitative Spatio-Temporal Reasoning

- Every day interaction with the physical world, **spatial reasoning is driven by qualitative abstractions** rather than complete quantitative knowledge.

- Driven by cognitive approaches, there has been work on QSR about spatial change. **Qualitative Spatio-Temporal Reasoning** encompasses all such techniques.

Qualitative Spatio-Temporal Reasoning

- One of the most **influential formalism** of qualitative representation of space is **Region Connection Calculus** (RCC), based on first-order logic.

- Contrary to traditional systems dealing with space, **basic entities** in Region Connection Calculus is not points but **non-empty regions of space**!

Region Connection Calculus

- RCC is a **mereotopological theory** i.e., defines property of regions of a topological space
  - starting from a single **binary topological relation** of connection, C(x, y), which is true if regions x and y are connected.
  - The topological primitive is used to define the **mereological relation** of parthood, P(x, y), which is true if x is a part of y.

Ontological Assumptions

- Primitive entities of the theory are to be interpreted as non-empty regular regions of space-time.

  4-D Regions: **Space-Time Histories**

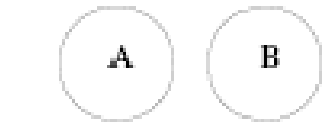
- Do not admit lower dimensional entities such as temporal points into our ontology.

  Pure pointless **Mereotopology**

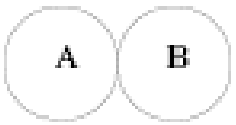
Region Connection Calculus

- The full **first-order theory** of RCC is too expressive to be computationally useful and is in fact undecidable.
- A propositional fragment of RCC, limited to **eight JEPD binary predicates** often referred to as RCC-8 is decidable.
- The language of RCC-8 consists of region variables and the eight JEPD binary relations

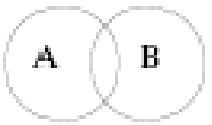
Region Connection Calculus



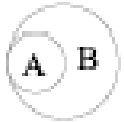
DC(A,B)
'A disconnected from B'



EC(A,B)
'A externally connected to B"

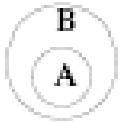


PO(A,B)
'A partially overlapping B'



TPP(A,B)
'A tangential proper part of B'

TPPi(B,A)
'B has tangential proper part A'



NTPP(A,B)
'A nontangential proper-part of B'

NTPPi(B,A)
'B has nontangential proper-part A'

EQUALS(A,B)
'A equal to B'

RCC-8: Eight JEPD Binary Relations

Skip Axiomatization

Axiomatisation of C_α

A1. $\mathsf{C}_\alpha(x, x)$

A2. $\mathsf{C}_\alpha(x, y) \rightarrow \mathsf{C}_\alpha(y, x)$

A3. $\forall z[\mathsf{C}_\alpha(z, x) \leftrightarrow \mathsf{C}_\alpha(z, y)] \rightarrow [x =_\alpha y]$

A4. $\exists z \forall u[\mathsf{C}_{\mathsf{st}}(u, z) \leftrightarrow (\mathsf{C}_{\mathsf{st}}(u, x) \vee \mathsf{C}_{\mathsf{st}}(u, y))]$

A5. $\neg \mathsf{P}_{\mathsf{st}}(x, y) \rightarrow \exists z \forall w[(\mathsf{P}_{\mathsf{st}}(w, x) \wedge \mathsf{DR}_{\mathsf{st}}(w, y)) \leftrightarrow \mathsf{P}_{\mathsf{st}}(w, z)]$

A6. $\mathsf{O}_{\mathsf{st}}(x, y) \rightarrow \exists z \forall u[\mathsf{C}_{\mathsf{st}}(u, z) \leftrightarrow \exists v(\mathsf{P}_{\mathsf{st}}(v, x) \wedge \mathsf{P}_{\mathsf{st}}(v, y) \wedge \mathsf{C}_{\mathsf{st}}(v, u))]$

A7. $\forall x[\exists y[\neg \mathsf{C}_{\mathsf{st}}(x, y) \rightarrow \exists z[\forall w(\mathsf{C}_{\mathsf{st}}(w, z) \leftrightarrow \neg \mathsf{NTPP}_{\mathsf{st}}(w, x)) \wedge \forall w(\mathsf{O}_{\mathsf{st}}(w, z) \leftrightarrow \neg \mathsf{P}_{\mathsf{st}}(w, x))]]]$

A8. $\mathsf{DR}_{\mathsf{st}}(x, y) \leftrightarrow \mathsf{Null}(x \cap y)$

Further Axioms for C_t

A9. $x \not\asymp y \rightarrow \neg(x <_t y)$

A10. $x <_t y \rightarrow \neg(y <_t x)$

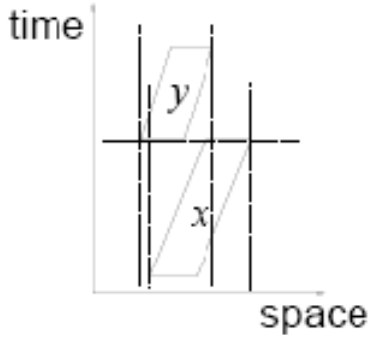
A11. $[x <_t y \wedge y \not\asymp z \wedge z <_t w] \rightarrow (x <_t w)$

A12. $x <_t y \rightarrow \forall z[(z \subseteq_t x \rightarrow z <_t y) \wedge (z \subseteq_t y \rightarrow x <_t z)]$

A13. $[x <_t y \wedge z <_t y] \leftrightarrow (x \cup z) <_t y$

A14. $(x \cup y) \not\asymp z \leftrightarrow [x \not\asymp z \vee y \not\asymp z]$

Further Axioms involving C_α



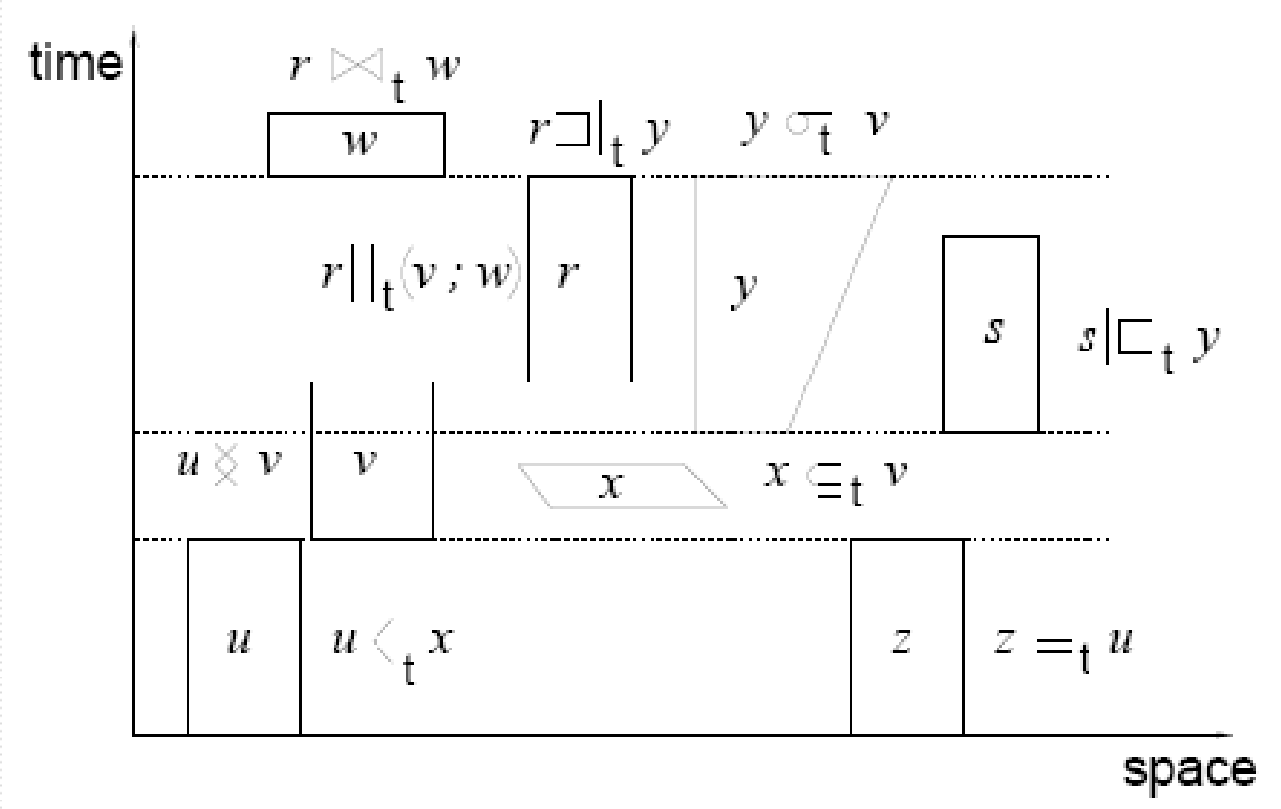
A15. $\mathsf{C}_{\mathsf{st}}(x,y) \rightarrow [\mathsf{C}_{\mathsf{t}}(x,y) \wedge \mathsf{C}_{\mathsf{sp}}(x,y)]$

A16. $\exists x \exists y [\mathsf{C}_{\mathsf{t}}(x,y) \wedge \neg \mathsf{C}_{\mathsf{st}}(x,y)]$

A17. $\exists x \exists y [\mathsf{C}_{\mathsf{sp}}(x,y) \wedge \neg \mathsf{C}_{\mathsf{st}}(x,y)]$

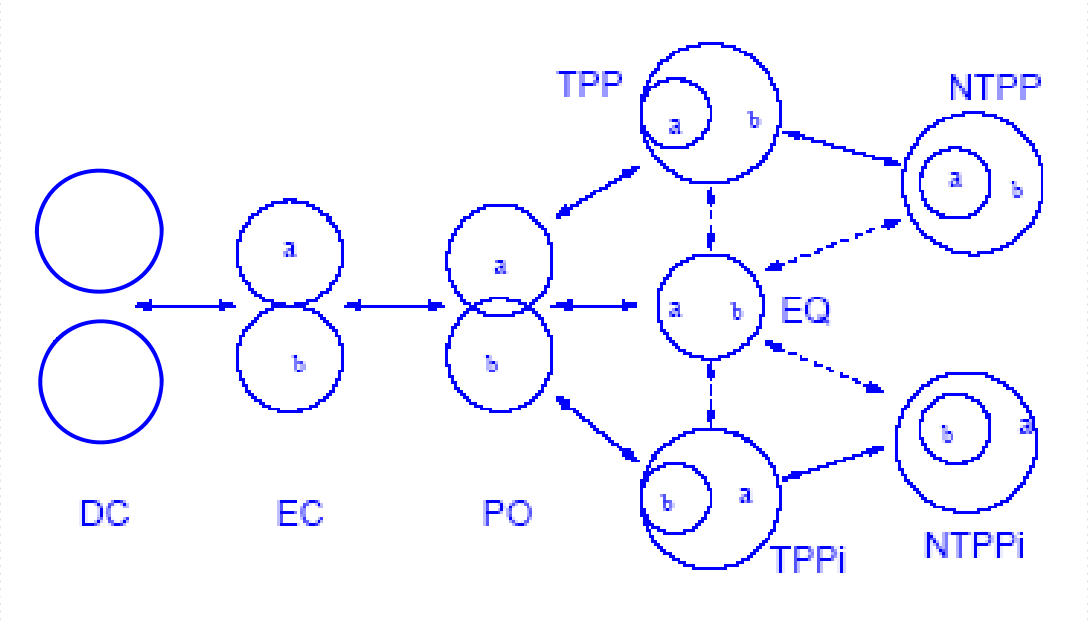
A18. $w \subseteq_{\mathsf{t}} y \rightarrow \exists x [\mathsf{TS}(x,y) \wedge x =_{\mathsf{t}} w]$

Region Connection Calculus



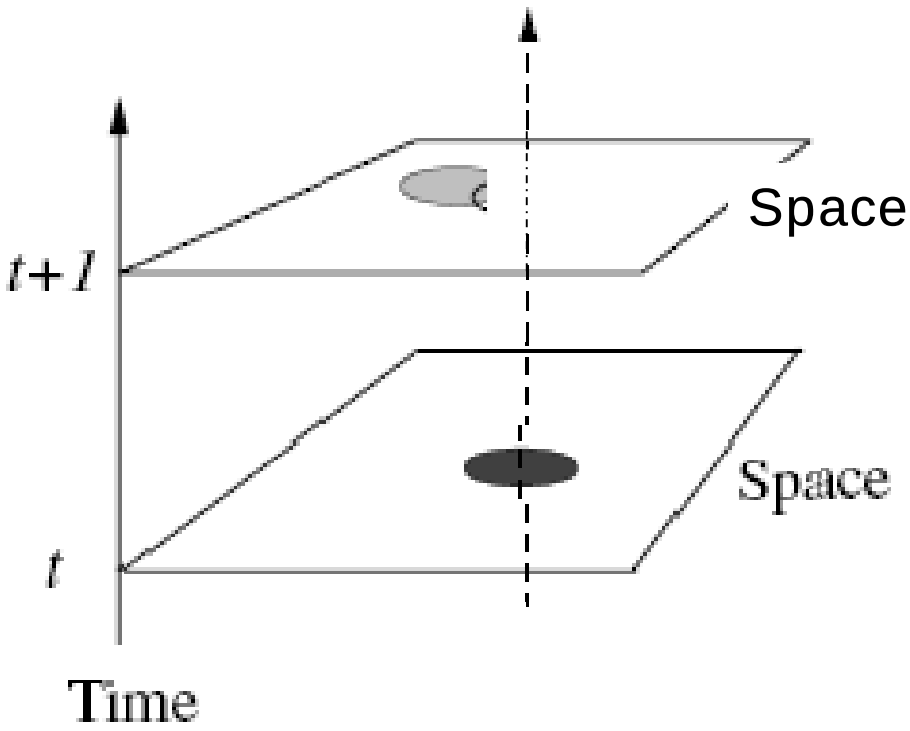
Temporal Relations over spatio-temporal regions

Conceptual Neighbourhood



Transition graph for Region Connection Calculus

Spatio-temporal Universe



Spatio-temporal Universe

- We can model **space as a quasi-order** $\mathcal{Q}$ and if the **flow of time** is represented by a **linear order** $\mathcal{T}$, then the whole spatio-temporal 'universe' can be viewed as **the Cartesian product** $\mathcal{Q} \times \mathcal{T}$, in which
  - The modalities act 'horizontally' to talk about spatial regions,
  - while the temporal operators act 'vertically' taking care of their movements in time.

Spatio-temporal Logic ST_0

- Combining RCC-8 with the PTL; capable of describing topological relationships that change over time.
  - Time is assumed to be isomorphic with the set of natural numbers.
  - Temporal ordering $<_t$ is defined together with two temporal operators - *Since* and *Until*.
- Application of *Since* and *Until* or other standard operators like $\bigcirc$ (next), $\Diamond$ (sometimes) or $\Box$ (always) to spatial formulae leads to ST_0.

Rule-based Filter in ST_0

- **Many Sorted Modal logic** is used in order to **handle the number of fields** involved in the condition of a rule. The sorted logic allow a **hierarchical sort structure** with sorts for rules R and individual sorts F_i for packet fields.

- Language contains **binary relation ≤ one for each sort**; expressing the **sequential order on rules for sort R** and **orders on the field's elements for each sort F_i**.

Rule-based Filter in ST_0

- We **consider the 'space' of all conditions**; and such **'carved-up' points are in essence sets**.
- Two such 'carved-up' points can be **related by one of the set-theoretic relations** $\{\subset, \supset, =\}$.
  - $C_{\subset}(x, y)$: x is '**subset connected**' to y; x and y are regions, with x's closure a subset of the closure of y.
  - $C_{=}(x, y)$: x is '**equal connected**' to y; x and y are regions whose closures are equal.
  - Connection primitive, $C_{\supset}(x, y)$: x is '**superset connected**' to y, is true just in case the closure of x is a superset of the closure of y.

Rule-based Filter in ST_0

- For the **mereological part of the spatio-temporal language ST_0**, we explicitly introduce the notion of parthood
  - P(x, y) : x is part of y.
- Include a **Boolean predicate d(R) to represent the 'action' of application of rule R**; with the intended interpretation that the condition is
  - accept (if true) or
  - deny (if false).

Firewall Rule Relations

- In order to build a reliable model for filtering rules, one need to **enumerate all possible relations** that may exist between two filtering rules.
- This has been addressed in the literature; and a **set of five relations have been identified**.
- For a pair of rules R1 and R2, these relations have been shown to be
  - **distinct**, i.e. only one relation can relate R1 and R2; and
  - **complete**, i.e. no other relation between R1 and R2 could exist.

Rule Relations

Definition 1 *Rules* R_1 *and* R_2 *are* completely disjoint (DR) *if every field in* R_1 *is not a subset nor a superset nor equal to the corresponding field in* R_2.

Definition 2 *Rules* R_1 *and* R_2 *are* exactly matching (EQ) *if every field in* R_1 *is equal to the corresponding field in* R_2.

Definition 3 *Rules* R_1 *and* R_2 *are* inclusively matching (PP) *if they do not exactly match and if every field in* R_1 *is a subset or equal to the corresponding field in* R_2.

Definition 4 *Rules* R_1 *and* R_2 *are* partially disjoint (PO) *if there is at least one field in* R_1 *that is a subset or a superset or equal to the corresponding field in* R_2*; and there is at least one field in* R_1 *that is a not a subset not a superset or not equal to the corresponding field in* R_2;.

Definition 5 *Rules* R_1 *and* R_2 *are* correlated (CR) *if some fields in* R_1 *are subsets or equal to the corresponding field in* R_2*; and the rest of the fields in* R_1 *are supersets of the corresponding field in* R_2.

Rule Relations - ST_0 Definable

$$\text{D1.}\quad \mathrm{DR}(\mathrm{R}_1,\mathrm{R}_2) \equiv_{def} \neg[\mathrm{C}_{\supset}(\mathrm{R}_1,\mathrm{R}_2) \wedge \mathrm{C}_{\subset}(\mathrm{R}_1,\mathrm{R}_2) \wedge \mathrm{C}_{=}(\mathrm{R}_1,\mathrm{R}_2)]$$

$$\text{D2.}\quad \mathrm{EQ}(\mathrm{R}_1,\mathrm{R}_2) \equiv_{def} \forall^{\mathrm{Fi}} p_1[\mathrm{P}(p_1,\mathrm{R}_1) \leftrightarrow \mathrm{P}(p_1,\mathrm{R}_2)]$$

$$\text{D3.}\quad \mathrm{PP}(\mathrm{R}_1,\mathrm{R}_2) \equiv_{def} \neg\mathrm{EQ}(\mathrm{R}_1,\mathrm{R}_2) \wedge \forall^{\mathrm{Fi}} p_1,p_2[[\mathrm{P}(p_1,\mathrm{R}_1) \wedge \mathrm{P}(p_2,\mathrm{R}_2)] \rightarrow [\mathrm{C}_{\subset}(p_1,p_2) \vee \mathrm{C}_{\supset}(p_1,p_2)]]$$

$$\text{D4.}\quad \mathrm{PO}(\mathrm{R}_1,\mathrm{R}_2) \equiv_{def} \exists^{\mathrm{Fi}} p_1,p_2[\mathrm{P}(p_1,\mathrm{R}_1) \wedge \mathrm{P}(p_2,\mathrm{R}_2) \wedge \mathrm{C}_{\alpha}(p_1,p_2)] \wedge \exists p_1,p_2[\mathrm{P}(p_1,\mathrm{R}_1) \wedge \mathrm{P}(p_2,\mathrm{R}_2) \wedge \neg\mathrm{C}_{\alpha}(p_1,p_2)]$$

$$\text{D5.}\quad \mathrm{CR}(\mathrm{R}_1,\mathrm{R}_2) \equiv_{def} \exists^{\mathrm{Fi}} p_1,p_2[\mathrm{P}(p_1,\mathrm{R}_1) \wedge \mathrm{P}(p_2,\mathrm{R}_2) \wedge [\mathrm{C}_{\subset}(p_1,p_2) \vee \mathrm{C}_{=}(p_1,p_2)]] \wedge \exists^{\mathrm{U-Fi}} p_1,p_2[\mathrm{P}(p_1,\mathrm{R}_1) \wedge \mathrm{P}(p_2,\mathrm{R}_2) \wedge \neg\mathrm{C}_{\supset}(p_1,p_2)]$$

Firewall rule relations can be expressed as ST_0 formula

Intra-Firewall Anomalies

Definition 6 *Shadowing anomaly: A rule* R_1 *is* shadowed *if there is a rule* R_2*, preceding* R_1 *in the filter, such that all packets that match* R_1 *already match* R_2*.*

Definition 7 *Redundancy anomaly: A redundant rule* R_1 *performs the same action on the same packets as another rule; therefore if the redundant rule is removed, the security policy will not be affected.*

Definition 8 *Correlation anomaly: Rules* R_1 *and* R_2 *have a correlation anomaly if* R_1 *and* R_2 *are correlated but their action is different.*

Definition 9 *Generalization anomaly: A rule* R_1 *is a generalization of a preceding rule* R_2 *if the later rule* inclusively matches (PP) *the preceding rule.*

Definition 10 *Irrelevance anomaly: A filtering rule* R_1 *in a firewall is irrelevant if this rule cannot match any traffic that might flow through this firewall. For irrelevance, the traffic through this firewall, formed of all packets which weren't rejected by upstream firewalls, must be expressed by an* ST_0 *formula TTF.*

Anomalies - ST_0 Definable

D6. Shadowing Anomaly: $\Box^{+}[EQ(R_1, R_2) \vee PP(R_1, R_2)] \wedge [R_2 < R_1]$

D7. Redundancy Anomaly: $\Box^{+}[EQ(R_1, R_2) \vee PP(R_1, R_2)] \wedge [R_2 < R_1] \wedge [d(R_1) = d(R_2)]$

D8. Correlation Anomaly: $\Diamond^{+}[CR(R_1, R_2)] \wedge [d(R_1) \neq d(R_2)]$

D9. Generalization Anomaly: $\Diamond^{+}[PP(R_1, R_2)] \wedge [d(R_1) \neq d(R_2)] \wedge [R_2 < R_1]$

D10. Irrelevance Anomaly: $[R_1 \cap TTF = \emptyset]$

Intra-firewall anomalies are ST_0 definable

Anomaly Detection - Model Checking

- **Anomaly detection** within such a framework is the **model-checking problem of a ST_0-formula**.
- If we consider the flow of time (N,<), then **satisfiability** in ST_0 **is PSPACE-complete**.
  - anomaly detection here is not more expensive than the existing algorithms.
  - existing algorithms account for only subset of anomalies, model-checking covers all anomalies at once.

Many Thanks!