# Spot's Temporal Logic Formulas

Alexandre Duret-Lutz <adl@lrde.epita.fr>compiled on February 23, 2023, for Spot 2.10.4

1.		·	2
		1	2
	1.2.	Usage in Model Checking	2
2.	Tem	nporal Syntax & Semantics	3
			3
			3
	2.2.		3
		1	4
			4
	2.3.		4
			5
			5
	2.4.	Temporal Operators	5
		* *	6
			6
			6
	25		7
	2.0.		7
			8
			9
	26	SERE-LTL Binding Operators	
	2.0.	2.6.1. Semantics	
		2.6.2. Syntactic Sugar	
		2.6.3. Trivial Identities (Occur Automatically)	
		2.0.5. Invial identifies (Occur Automatically)	1
3.	Grai	mmar 1	2
	3.1.	Operator precedence	2
4.		perties 1	
		Pure Eventualities and Purely Universal Formulas	
	4.2.	Syntactic Hierarchy Classes	5
_	D		_
Э.		vritings 1	
		Unabbreviations	
		LTL simplifier	
		Negative normal form	
	5.4.		
		5.4.1. Basic Simplifications	
		5.4.2. Simplifications for Eventual and Universal Formulas	
		5.4.3. Simplifications Based on Implications	4

A. Defining LTL with only one of U, $W$ , $R$ , or $M$	27
B. Syntactic Implications	29
Bibliography	29

# 1. Reasoning with Infinite Sequences

### 1.1. Finite and Infinite Sequences

Let  $N = \{0, 1, 2, ...\}$  denote the set of natural numbers and  $\omega \notin N$  the first transfinite ordinal. We extend the < relation from N to  $N \cup \{\omega\}$  with  $\forall n \in N, n < \omega$ . Similarly let us extend the addition and subtraction with  $\forall n \in N, \omega + n = \omega - n = \omega + \omega = \omega$ .

For any set A, and any number  $n \in N \cup \{\omega\}$ , a sequence of length n is a function  $\sigma : \{0, 1, \ldots, n-1\} \to A$  that associates each index i < n to an element  $\sigma(i) \in A$ . The sequence of length 0 is a particular sequence called the *empty word* and denoted  $\varepsilon$ . We denote  $A^n$  the set of all sequences of length n on A (in particular  $A^\omega$  is the set of infinite sequences on A), and  $A^* = \bigcup_{n \in N} A^n$  denotes the set of all finite sequences. The length of any sequence  $\sigma$  is noted  $|\sigma|$ , with  $|\sigma| \in N \cup \{\omega\}$ .

For any sequence  $\sigma$ , we denote  $\sigma^{i...j}$  the finite subsequence built using letters from  $\sigma(i)$  to  $\sigma(j)$ . If  $\sigma$  is infinite, we denote  $\sigma^{i...}$  the suffix of  $\sigma$  starting at letter  $\sigma(i)$ .

### 1.2. Usage in Model Checking

The temporal formulas described in this document, should be interpreted on behaviors (or executions, or scenarios) of the system to verify. In model checking we want to ensure that a formula (the property to verify) holds on all possible behaviors of the system.

If we model the system as some sort of giant automaton (e.g., a Kripke structure) where each state represent a configuration of the system, a behavior of the system can be represented by an infinite sequence of configurations. Each configuration can be described by an affectation of some proposition variables that we will call *atomic propositions*. For instance r = 1, y = 0, g = 0 describes the configuration of a traffic light with only the red light turned on.

Let AP be a set of atomic propositions, for instance AP =  $\{r, y, g\}$ . A configuration of the model is a function  $\rho: AP \to B$  (or  $\rho \in B^{AP}$ ) that associates a truth value ( $B = \{0, 1\}$ ) to each atomic proposition. A behavior of the model is an infinite sequence  $\sigma$  of such configurations. In other words:  $\sigma \in (B^{AP})^{\omega}$ .

When a formula  $\varphi$  holds on an *infinite* sequence  $\sigma$ , we write  $\sigma \varphi$  (read as  $\sigma$  is a model of  $\varphi$ ). When a formula  $\varphi$  holds on an *finite* sequence  $\sigma$ , we write  $\sigma \varphi$ .

# 2. Temporal Syntax & Semantics

### 2.1. Boolean Constants

The two Boolean constants are '1' and '0'. They can also be input as 'true' or 'false' (case insensitive) for compatibility with the output of other tools, but Spot will always use '1' and '0' in its output.

#### 2.1.1. Semantics

 $\sigma \mathsf{0}$ 

 $\sigma$ 1

### 2.2. Atomic Propositions

Atomic propositions in Spot are strings of characters. There are no restrictions on the characters that appear in the strings, however because some of the characters may also be used to denote operators you may have to represent the strings differently if they include these characters.

- 1. Any string of characters represented between double quotes is an atomic proposition.
- 2. Any sequence of alphanumeric characters (including '\_') that is not a reserved keyword and that starts with a characters that is not an uppercase 'F', 'G', or 'X', is also an atomic proposition. In this case the double quotes are not necessary.
- 3. Any sequence of alphanumeric character that starts with 'F', 'G', or 'X', has a digit in second position, and anything afterwards, is also an atomic propositions, and the double quotes are not necessary.

Here is the list of reserved keywords:

- 'true', 'false' (both are case insensitive)
- 'F', 'G', 'M', 'R', 'U', 'V', 'W', 'X', 'xor'

The only way to use an atomic proposition that has the name of a reserved keyword, or one that starts with a digit, is to use double quotes.

The reason we deal with leading 'F', 'G', and 'X' specifically in rule 2 is that these are unary LTL operators and we want to be able to write compact formulas like 'GFa' instead of the equivalent 'G(F(a))' or 'GF(a)'. If you want to name an atomic proposition 'GF(a)', you will have to quote it as '"GF(a)".

The exception done by rule 3 when these letters are followed by a digit is meant to allow 'X0', 'X1', 'X2', ... to be used as atomic propositions. With only rule 2, 'X0' would be interpreted as 'X(0)', that is, the LTL operator X applied to the constant *false*, but there is really little reason to use such a construction in a formula (the same is true for 'F' and 'G', and also when applied to '1'). On the other hand, having numbered versions of a variable is pretty common, so it makes sense to favor this interpretation.

If you are typing in formulas by hand, we suggest you name all your atomic propositions in lower case, to avoid clashes with the uppercase operators.

If you are writing a tool that produces formula that will be feed to Spot and if you cannot control the atomic propositions that will be used, we suggest that you always output atomic propositions between double quotes to avoid any unintended misinterpretation.

### 2.2.1. Examples

- '"a<=b+c"' is an atomic proposition. Double quotes can therefore be used to embed constructs specific to the underlying formalism, and still regard the resulting construction as an atomic proposition.
- 'light\_on' is an atomic proposition.
- 'Fab' is not an atomic proposition, this is actually equivalent to the formula 'F(ab)' where the temporal operator F is applied to the atomic proposition 'ab'.
- 'FINISHED' is not an atomic proposition for the same reason; it actually stands for 'F(INISHED)'
- 'F100ZX' is an atomic proposition by rule 3.
- 'FX100' is not an atomic proposition, it is equivalent to the formula 'F(X100)', where 'X100' is an atomic proposition by rule 3.

#### 2.2.2. Semantics

For any atomic proposition a, we have

$$\sigma a \iff \sigma(0)(a) = 1$$

In other words a holds if and only if it is true in the first configuration of  $\sigma$ .

### 2.3. Boolean Operators (for Temporal Formulas)

Two temporal formulas f and g can be combined using the following Boolean operators:

operation	preferred syntax	other supported syntaxes			UTF8 char- preferred	acters supported others
negation	! <i>f</i>	~ <i>f</i>	C \ /	<i>C</i> .	¬ U+00AC	
disjunction	$f \mid g$	$f \mid \mid g$	f \/ g	f + g	∨ U+2228	∪ U+222A
conjunction	f & g	f && g	$f \wedge g$	$f * g^1$	∧ U+2227	∩ U+2229
implication	f -> g	f => g	f> g		ightarrow U+2192	$ ightarrow$ U+27F6, $\Rightarrow$ U+21D2 U+27F9
exclusion	f xor $g$	f ^ g			$\oplus$ U+2295	
equivalence	f <-> g	f <=> g	f <> g		$\leftrightarrow \texttt{U+2194}$	⇔ U+21D4

Additionally, an atomic proposition a can be negated using the syntax 'a=0', which is equivalent to '! a'. Also 'a=1' is equivalent to just 'a'. These two syntaxes help us read formulas written using Wring's syntax.

When using UTF-8 input, a '=0' that follow a single-letter atomic proposition may be replaced by a combining overline U+0305 or a combining overbar U+0304. When instructed to emit UTF-8, Spot will output ' $\bar{a}$ ' using a combining overline instead of ' $\neg a$ ' for any single-letter atomic proposition.

When a formula is built using only Boolean constants (section 2.1), atomic proposition (section 2.2), and the above operators, we say that the formula is a *Boolean formula*.

<sup>&</sup>lt;sup>1</sup>The \*-form of the conjunction operator (allowing better compatibility with Wring and VIS) may only used in temporal formulas. Boolean expressions that occur inside SERE (see Section 2.5) may not use this form because the \* symbol is used as the Kleen star.

### 2.3.1. Semantics

$$\sigma ! f \iff (\sigma f)$$

$$\sigma f \& g \iff (\sigma f) \land (\sigma g)$$

$$\sigma f \mid g \iff (\sigma f) \lor (\sigma g)$$

$$\sigma f \rightarrow g \iff (\sigma f) \lor (\sigma g)$$

$$\sigma f \text{xor } g \iff ((\sigma f) \land (\sigma g)) \lor ((\sigma f) \land (\sigma g))$$

$$\sigma f < \Rightarrow g \iff ((\sigma f) \land (\sigma g)) \lor ((\sigma f) \land (\sigma g))$$

### 2.3.2. Trivial Identities (Occur Automatically)

Trivial identities are applied every time an expression is constructed. This means for instance that there is not way to construct the expression '!! a' in Spot, such an attempt will always yield the expression 'a'.

$$\begin{array}{ll} !\ 0\equiv 1 & 1\rightarrow f\equiv f & f\rightarrow 1\equiv 1 \\ !\ 1\equiv 0 & 0\rightarrow f\equiv 1 & f\rightarrow 0\equiv !\ f \\ !\ !\ f\equiv f & f\rightarrow f\equiv 1 \end{array}$$

The next set of rules apply to operators that are commutative, so these identities are also valid with the two arguments swapped.

$$\begin{array}{lll} 0 \& f \equiv 0 & 0 & | f \equiv f & 0 & \text{for } f \equiv f \\ 1 \& f \equiv f & 1 & | f \equiv 1 & 1 & \text{for } f \equiv ! & f \\ f \& f \equiv f & f & | f \equiv f & f & \text{for } f \equiv 0 & f < - > f \equiv 1 \end{array}$$

The '&' and '|' operators are associative, so they are actually implemented as *n*-ary operators (i.e., not binary): this allows us to reorder all arguments in a unique way (e.g. alphabetically). For instance the two expressions 'a&c&b&!d' and 'c&!d&b&a' are actually represented as the operator '&' applied to the arguments {a,b,c,!d}. Because these two expressions have the same internal representation, they are actually considered equal for the purpose of the above identities. For instance '(a&c&b&!d)->(c&!d&b&a)' will be rewritten to '1' automatically.

### 2.4. Temporal Operators

Given two temporal formulas f, and g, the following temporal operators can be used to construct another temporal formula.

	preferred	other supported	upported UTF8 characters sup	
operator	syntax	syntaxes	preferred	others
(Weak) Next	X f	() f	○ U+25CB	○ U+25EF
Strong Next	X[!]f			
Eventually	F f	<> f	♦ U+25C7	♦ U+22C4 U+2662
Always	$\mathbf{G} f$	[]f	□ U+25A1	☐ U+2B1C U+25FB
(Strong) Until	f U g			
Weak Until	f W g			
(Weak) Release	f R g	$f \vee g$		
Strong Release	f M g			

#### 2.4.1. Semantics

$$\sigma X f \iff \sigma^{1...} f 
\sigma X [!] f \iff \sigma^{1...} f 
\sigma F f \iff \exists i \in N, \sigma^{i...} f 
\sigma G f \iff \forall i \in N, \sigma^{i...} f 
\sigma f U g \iff \exists j \in N, \begin{cases} \forall i < j, \sigma^{i...} f \\ \sigma^{j...} g \end{cases} 
\sigma f W g \iff (\sigma f U g) \lor (\sigma G f) 
\sigma f M g \iff \exists j \in N, \begin{cases} \forall i \le j, \sigma^{i...} g \\ \sigma^{j...} f \end{cases} 
\sigma f R g \iff (\sigma f M g) \lor (\sigma G g)$$

Note that the semantics of X (weak next) and X[!] (strong next) are identical in LTL formulas. The two operators make sense only to build LTLf formulas (i.e., LTL with finite semantics), for which support is being progressively introduced in Spot.

Appendix A explains how to rewrite the above LTL operators using only X and one operator chosen among U, W, M, and R. This could be useful to understand the operators R, M, and W if you are only familiar with X and U.

### 2.4.2. Syntactic Sugar

The syntax on the left is equivalent to the syntax on the right. Some of rewritings taken from the syntax of TSLF [?] are performed from left to right when parsing a formula. They express the fact that some formula f has to be true in n steps, or at some or all times between n and m steps.

$$X[n]f \equiv \underbrace{XX...X}f$$

$$n \text{ occurrences of } X$$

$$F[n:m]f \equiv \underbrace{XX...X}(f \mid \underbrace{X(f \mid X(... \mid Xf))})$$

$$n \text{ occ. of } X$$

$$m-n \text{ occ. of } X$$

$$X[n:m]f \equiv \underbrace{XX...X}(f \& \underbrace{X(f \& X(... \& Xf))})$$

$$n \text{ occ. of } X$$

$$X[n:m]f \equiv \underbrace{X[n]f} = \underbrace{X[$$

### 2.4.3. Trivial Identities (Occur Automatically)

$$X[!]0 \equiv 0$$
  $F0 \equiv 0$   $G0 \equiv 0$   $X1 \equiv 1$   $Ff f \equiv Ff$   $GG f \equiv Gf$ 

$f$ U 1 $\equiv$ 1	$f$ W 1 $\equiv$ 1	$f{ m M}{ m O}\equiv{ m O}$	$f$ R 1 $\equiv$ 1
$\mathtt{O}\mathtt{U}f\equiv f$	$\mathtt{O}\mathtt{W}f\equiv f$	${\tt OM}f\equiv{\tt O}$	$f$ R O $\equiv$ O
$f$ U O $\equiv$ O	1 W $f\equiv$ 1	$\mathtt{1}\mathtt{M}f\equiv f$	1 R $f \equiv f$
$f$ U $f \equiv f$	$f  \mathtt{W}  f \equiv f$	$f$ M $f \equiv f$	$f R f \equiv f$

### 2.5. SERE Operators

The "SERE" acronym will be translated to different word depending on the source. It can mean either "Sequential Extended Regular Expression" [? ], "Sugar Extended Regular Expression" [? ], or "Semi-Extended Regular Expression" [? ]. In any case, the intent is the same: regular expressions with traditional operations (union '|', concatenation ';', Kleen star '[\*]') are extended with operators such as intersection '&&', and fusion ':'.

Any Boolean formula (section 2.3) is a SERE. SERE can be further combined with the following operators, where f and g denote arbitrary SERE.

operation	preferred syntax		supported ntaxes		UTF8 charapreferred	acters supported others
empty word union	[*0] f   g	$f \mid \mid g$	$f \lor /g$	f + g		V U+2228 ∪ U+222A
intersection	f && g	$f \wedge g$			∩ U+2229	∧ U+2227
NLM intersection <sup>2</sup>	f & g					
concatenation	f; $g$					
fusion	f:g					
bounded ;-iter.	f [*ij]	f [*i:j]	f[*i  to  j]	f[*i,j]		
unbounded ;-iter.	f[*i]	f[*i:]	f[*i to]	f[*i,]		
bounded :-iter.	f[:*ij]	f [:*i:j]	f[:*i  to  j]	f[:*i,j]		
unbounded :-iter.	f[:*i]	f[:*i:]	f[:*i to]	f[:*i,]		
first match f	irst_match(	f)				

The character '\$' or the string 'inf' can also be used as value for j in the above operators to denote an unbounded range.<sup>3</sup> For instance 'a[\*i,\$]', 'a[\*i:inf]' and 'a[\*i.]' all represent the same SERE.

#### 2.5.1. Semantics

The following semantics assume that f and g are two SEREs, while a is an atomic proposition.

$$\begin{array}{c} \sigma 0 \\ \sigma 1 \iff |\sigma| = 1 \\ \sigma [*0] \iff |\sigma| = 0 \\ \sigma a \iff \sigma (0)(a) = 1 \wedge |\sigma| = 1 \\ \sigma f \mid g \iff (\sigma f) \vee (\sigma g) \\ \sigma f \&\& g \iff (\sigma f) \wedge (\sigma g) \\ \sigma f \& g \iff \exists k \in N, \begin{cases} \text{either } (\sigma f) \wedge (\sigma^{0.k-1} g) \\ \text{or } (\sigma^{0.k-1} f) \wedge (\sigma g) \end{cases} \\ \sigma f ; g \iff \exists k \in N, (\sigma^{0.k-1} f) \wedge (\sigma^{k} g) \end{array}$$

 $<sup>^2</sup>Non$ -Length-Matching interesction.

<sup>&</sup>lt;sup>3</sup>SVA uses '\$' while PSL uses 'inf'.

$$\sigma f: g \iff \exists k \in N, (\sigma^{0..k}f) \land (\sigma^{k..}g)$$
 
$$\sigma f[*i..j] \iff \begin{cases} \text{either} & i = 0 \land \sigma = \varepsilon \\ \text{or} & i = 0 \land j > 0 \land (\exists k \in N, (\sigma^{0..k-1}f) \land (\sigma^{k..}f[*0..j-1])) \\ \text{or} & i > 0 \land j > 0 \land (\exists k \in N, (\sigma^{0..k-1}f) \land (\sigma^{k..}f[*i-1..j-1])) \end{cases}$$
 
$$\sigma f[*i..] \iff \begin{cases} \text{either} & i = 0 \land \sigma = \varepsilon \\ \text{or} & i = 0 \land (\exists k \in N, (\sigma^{0..k-1}f) \land (\sigma^{k..}f[*i-1..])) \\ \text{or} & i > 0 \land (\exists k \in N, (\sigma^{0..k-1}f) \land (\sigma^{k..}f[*i-1..])) \end{cases}$$
 
$$\sigma f[:*i..j] \iff \begin{cases} \text{either} & i = 0 \land j = 0 \land \sigma 1 \\ \text{or} & i = 0 \land j > 0 \land (\exists k \in N, (\sigma^{0..k}f) \land (\sigma^{k..}f[:*i-1..j-1])) \\ \text{or} & i > 0 \land j > 0 \land (\exists k \in N, (\sigma^{0..k}f) \land (\sigma^{k..}f[:*i-1..j-1])) \end{cases}$$
 
$$\sigma f[:*i..] \iff \begin{cases} \text{either} & i = 0 \land \sigma 1 \\ \text{or} & i = 0 \land \sigma 1 \\ \text{or} & i = 0 \land (\exists k \in N, (\sigma^{0..k}f) \land (\sigma^{k..}f[:*i-1..j-1])) \\ \text{or} & i > 0 \land (\exists k \in N, (\sigma^{0..k}f) \land (\sigma^{k..}f[:*i-1..])) \end{cases}$$
 
$$\sigma f[:*i..] \iff \begin{cases} \sigma f \land (\forall k < |\sigma|, \sigma^{0..k}f) \land (\sigma^{k..}f[:*i-1..]) \end{cases}$$

Notes:

- The semantics of && and & coincide if both operands are Boolean formulas.
- The SERE f:g will never hold on [\*0], regardless of the value of f and g. For instance a[\*]:b[\*] is actually equivalent to  $a[*];\{a \&\& b\};b[*]$ .
- The [:\*i..] and [:\*i..j] operators are iterations of the : operator just like The [\*i..] and [\*i..j] are iterations of the ; operator. More graphically:

$$f[*i..j] = \underbrace{f; f; ...; f}_{\text{between } i \text{ and } j \text{ copies of } f}$$
 $f[:*i..j] = \underbrace{f: f: ...: f}_{\text{between } i \text{ and } j \text{ copies of } f}$ 

with the convention that

$$f[*0..0] = [*0]$$
  $f[:*0..0] = 1$ 

- The [:\*i..] and [:\*i..j] operators are not defined in PSL. While the bounded iteration can be seen as syntactic sugar on :, the unbounded version really is a new operator.
  - [:\*1..], for which we define the [:+] syntactic sugar below, actually corresponds to the  $^{\oplus}$  operator introduced by ? ]. With this simple addition, it is possible to define a subset of PSL that expresses exactly the stutter-invariant  $\omega$ -regular languages.
- The first\_match operator does not exist in PSL. It comes from SystemVerilog Assertions (SVA) [? ]. One intuition behind first\_match(f) is that the DFA for first\_match(f) can be obtained from the DFA for f by removing all transitions leaving accepting states.

#### 2.5.2. Syntactic Sugar

The syntax on the left is equivalent to the syntax on the right. These rewritings are performed from left to right when parsing a formula, and *some* are performed from right to left when writing it for output. *b* must be a Boolean formula.

```
\begin{array}{l} b \, [->i \, . \, .j] \, \equiv \, \{\{\, ! \, b\} \, [*o \, . \, ] \, ; \, b\} \, [*i \, . \, j] \, b \, [=i \, . \, j] \, \equiv \, \{\{\, ! \, b\} \, [*o \, . \, ] \, ; \, b\} \, [*i \, . \, .j] \, ; \, \{\, ! \, b\} \, [*o \, . \, ] \, ; \, b\} \, [*o \, . \, ] \, if \, i > 0 \\ b \, [=o \, . \, .] \, \equiv \, 1 \, [*o \, . \, .] \end{array}
```

$$f* \equiv f \, [*0 \, . \, ] \\ f[*] \equiv f \, [*0 \, . \,$$

 $[*k] \equiv 1[*k..k]$   $[*] \equiv 1[*0..]$   $[+] \equiv 1[*1..]$ 

The following adds input support for the SVA concatenation (or delay) operator [?]. The simplest equivalence are that f ##0 g, f ##1 g, f ##2 g mean respectively f: g, f; g, and f; 1; g, but the delay can be a range, and f can be omitted.

```
\begin{split} f \#\#[i..j] & g \equiv f \;;\; 1[*i-1..j-1] \;;\; g \quad \text{if } i > 0 \\ f \#\#[0..j] & g \equiv f \;:\; (1[*0..j] \;;\; g) \quad \text{if } \varepsilon f \\ f \#\#[0..j] & g \equiv (f \;;\; 1[*0..j]) \;:\; g) \quad \text{if } \varepsilon f \wedge \varepsilon g \\ f \#\#[0..j] & g \equiv (f \;;\; g) \mid (f \;;\; 1[*0..j-1] \;;\; g) \quad \text{if } \varepsilon f \wedge \varepsilon g \\ f \#\#[i..] & g \equiv f \;;\; 1[*i-1..] \;;\; g \quad \text{if } i > 0 \\ f \#\#[0..] & g \equiv f \;:\; (1[*] \;;\; g) \quad \text{if } \varepsilon f \\ f \#\#[0..] & g \equiv (f \;;\; 1[*]) \;:\; g) \quad \text{if } \varepsilon f \wedge \varepsilon g \\ f \#\#[0..] & g \equiv (f \;;\; g) \mid (f \;;\; 1[*] \;;\; g) \quad \text{if } \varepsilon f \wedge \varepsilon g \end{split}
```

#### 2.5.3. Trivial Identities (Occur Automatically)

The following identities also hold if j or l are missing (assuming they are then equal to  $\infty$ ). f can be any SERE, while b,  $b_1$ ,  $b_2$  are assumed to be Boolean formulas.

$$\begin{array}{c} \texttt{O}[*o..j] \equiv [*o] \\ \texttt{[*o]}[*i..j] \equiv [*o] \\ f[*i..j] [*k..l] \equiv f [*ik..jl] \text{ if } i(k+1) \leq jk+1 \\ f[*o] \equiv [*o] \\ f[*i..j] [*k..l] \equiv f \\ f[*il] \equiv f \\ f[*o].[*o..j] \equiv 1 \\ \texttt{[*o]}[:*o..j] \equiv 1 \\ f[*o].[:*o..j] \equiv 0 \text{ if } i > 0 \\ f[:*i..j] [:*k..l] \equiv f \text{ if } i > 0 \\ f[:*i..j] [:*k..l] \equiv f \text{ if } i > 0 \\ f[:*o] \equiv 1 \\ f[:*o] \equiv 1 \\ f[:*o] \equiv 1 \\ f[:*o] \equiv f \text{ if } i \in f \\ first_match(f) \equiv [*o] \text{ if } i \in f \\ first_match(f) \equiv f$$

The following rules are all valid with the two arguments swapped.

$$0 \& f \equiv 0 \qquad 0 \& \& f \equiv 0 \qquad 0 \mid f \equiv f \qquad 0 : f \equiv 0 \qquad 0 ; f \equiv 0$$
 
$$1 \& f \equiv \begin{cases} 1 \text{ if } \varepsilon f \\ f \text{ if } \varepsilon f \end{cases} \qquad 1 \& \& b \equiv b \qquad 1 \mid b \equiv 1 \qquad 1 : f \equiv f \text{ if } \varepsilon f \end{cases}$$
 
$$[*] \& f \equiv f \qquad [*] \mid f \equiv 1 [*] \qquad [*] ; f \equiv [*] \text{ if } \varepsilon f$$
 
$$[*0] \& f \equiv f \qquad [*0] \& \& f \equiv \begin{cases} [*0] \text{ if } \varepsilon f \\ 0 \text{ if } \varepsilon f \end{cases} \qquad [*0] : f \equiv 0 \qquad [*0] ; f \equiv f \end{cases}$$
 
$$f \& f \equiv f \qquad f \& \& f \equiv f \qquad f \mid f \equiv f \qquad f : f \equiv f [:*2] \qquad f ; f \equiv f [*2]$$
 
$$b_1 \& b_2 \equiv b_1 \& \& b_2 \qquad b_1 : b_2 \equiv b_1 \& \& b_2 \qquad f [*i..j] ; f [*k..l] \equiv f [*i+k..j+l]$$
 
$$f [:*i..j] : f [:*k..l] \equiv f [:*i+k..j+l]$$
 
$$f [:*i..j] : f [:*k..l] \equiv f [:*i+k..j+l]$$

### 2.6. SERE-LTL Binding Operators

The following operators combine a SERE *r* with a PSL formula *f* to form another PSL formula.

operation	preferred syntax	other supported syntaxes	
(universal) suffix implication existential suffix implication	{r}[]->f {r}<>->f	{ <i>r</i> }   -> <i>f</i>	{ <i>r</i> }( <i>f</i> )
weak closure negated weak closure	{r} !{r}		

For technical reasons, the negated weak closure is actually implemented as an operator, even if it is syntactically and semantically equal to the combination of ! and  $\{r\}$ .

UTF-8 input may combine one box or diamond character from section 2.4 with one arrow character from section 2.3 to replace the operators []->, <>->, as well as the operators []=> and <>=> that will be defined in 2.6.2. Additionally, |-> may be replaced by  $\mapsto$  U+21A6, and |=> by U+2907.

### 2.6.1. Semantics

The following semantics assume that r is a SERE, while f is a PSL formula.

$$\begin{split} \sigma\{r\} &<> > f \iff \exists k \geq 0, (\sigma^{0..k}r) \land (\sigma^{k..}f) \\ \sigma\{r\} &[ ] -> f \iff \forall k \geq 0, (\sigma^{0..k}r) \rightarrow (\sigma^{k..}f) \\ \sigma\{r\} \iff (\exists k \geq 0, \sigma^{0..k}r) \lor (\forall k \geq 0, \exists \pi \in (B^{AP})^*, (\sigma^{0..k} \prec \pi) \land (\pi r)) \\ \sigma! &\{r\} \iff (\forall k \geq 0, \sigma^{0..k}r) \land (\exists k \geq 0, \forall \pi \in (B^{AP})^*, (\sigma^{0..k} \prec \pi) \rightarrow (\pi r)) \end{split}$$

### 2.6.2. Syntactic Sugar

The syntax on the left is equivalent to the syntax on the right. These rewritings are performed from left to right when parsing a formula. Except the one marked with  $\stackrel{\dagger}{=}$ , the opposite rewritings are also preformed on output to ease reading.

$$\{r\} <>=> f \equiv \{r ; 1\} <>-> f$$
 
$$\{r\} [] => f \equiv \{r ; 1\} [] -> f$$
 
$$\{r\} ! \equiv \{r\} <>-> 1$$
 
$$\{r\} |=> f \stackrel{\dagger}{\equiv} \{r ; 1\} [] -> f$$

[]=> and |=> are synonyms in the same way as []-> and |-> are. The  $\{r\}$ ! operator is a *strong closure* operator.

### 2.6.3. Trivial Identities (Occur Automatically)

For any PSL formula *f* , any SERE *r* , and any Boolean formula *b* , the following rewritings are systematically performed (from left to right).

### 3. Grammar

For simplicity, this grammar gives only one rule for each operator, even if the operator has multiple synonyms (like '|', '||', and '\/').

```
constant ::= 0 \mid 1
atomic_prop ::= see secn 2.2
                                               (bformula)
                                                                           | bformula xor bformula
   bformula := constant
               | atomic_prop
                                              ! bformula
                                                                           | bformula <-> bformula
                                                                           | bformula -> bformula
                | atomic_prop=0
                                              | bformula & bformula
                                              | bformula | bformula
                | atomic_prop=1
        sere ::= bformula
                                              | [*i..j]
                                                                           ##i sere
               | { sere }
                                               [+]
                                                                           | ##[i..j] sere
                (sere)
                                              | sere [*i..j]
                                                                           | sere ##i sere
                | sere | sere
                                               | sere [+]
                                                                           | sere ## [i . . j] sere
                                               | sere[:*i..j]
                                                                           | first_match(sere)
                | sere & sere
                                              | sere[:+]
                sere && sere
                                               | sere [=i . . j]
                | sere ; sere
                                              | sere [->i . . j]
                | sere : sere
   tformula ::= bformula
                                              | X tformula
                                                                           | {sere} [] -> tformula
               (tformula)
                                              | X[!] tformula
                                                                           | {sere} [] => tformula
                ! tformula
                                              | X[i..j] t formula
                                                                           | {sere} <>-> tformula
                | tformula & tformula
                                              | X[i..j!] t formula
                                                                           | {sere} <>=> tformula
                | tformula | tformula
                                              | F tformula
                                                                           | {sere}
                | tformula -> tformula
                                              \mid F[i..j] t formula
                                                                           | {sere} !
                                              | F[i..j!] t formula
                | tformula xor tformula
                | tformula <-> tformula
                                              | G tformula
                | tformula U tformula
                                              |G[i..j]| tformula
                | tformula W tformula
                                              |G[i..j!]| tformula
                | tformula R tformula
                | tformula M tformula
```

### 3.1. Operator precedence

The following operator precedence describes the current parser of Spot. It has not always been this way. Especially, all operators were left associative until version 0.9, when we changed the associativity of ->, <->, U, R, W, and M to get closer to the PSL standard [??].

```
assoc.
        operators
                                                             priority
         []->, []=>, <>->, <>=>
right
                                                              lowest
 left
 left
 left
        ##i, ##[i . . j]
right
        ->, <->
 left
        xor
 left
 left
        &, &&
right
        U, W, M, R
        F, G, F[i..j], G[i..j]
        X, X[i..j]
         [*i..j], [+], [:*i..j], [:+], [=i..j], [->i..j]
         =0, =1
                                                              highest
```

Beware that not all tools agree on the associativity of these operators. For instance Spin, Itl2ba (same parser as spin), Wring, psl2ba, Modella, and NuSMV all have U and R as left-associative, while Goal (hence Büchi store), LTL2AUT, and LTL2Büchi (from JavaPathFinder) have U and R as right-associative. Vis and LBTT have these two operators as non-associative (parentheses required). Similarly the tools do not aggree on the associativity of -> and <->: some tools handle both operators as left-associative, or both right-associative, other have only -> as right-associative.

# 4. Properties

When Spot builds a formula (represented by an AST with shared subtrees) it computes a set of properties for each node. These properties can be queried from any spot::formula instance using the following methods:

is\_boolean() Whether the formula uses only Boolean operators. is\_sugar\_free\_boolean() Whether the formula uses only &, |, and ! operators. (Especially, no -> or <-> are allowed.) Whether the formula is in negative normal form. See section 5.3. is\_in\_nenoform() is\_X\_free() Whether the formula avoids the X operator. is\_ltl\_formula() Whether the formula uses only LTL operators. (Boolean operators are also allowed.) Whether the formula uses only PSL operators. (Boolean and LTL is\_psl\_formula() operators are also allowed.) is\_sere\_formula() Whether the formula uses only SERE operators. (Boolean operators are also allowed, provided no SERE operator is negated.) is\_finite() Whether a SERE describes a finite language (no unbounded stars), or an LTL formula uses no temporal operator but X. is\_eventual() Whether the formula is a pure eventuality. Whether the formula is purely universal. is\_universal() is\_syntactic\_safety() Whether the formula is a syntactic safety property. is\_syntactic\_guarantee() Whether the formula is a syntactic guarantee property. Whether the formula is a syntactic obligation property. is\_syntactic\_obligation() is\_syntactic\_recurrence() Whether the formula is a syntactic recurrence property. is\_syntactic\_persistence() Whether the formula is a syntactic persistence property. Whether the formula contains a special "marked" version of the is\_marked()  $\rightarrow$  or ! $\{r\}$  operators. Whether the formula accepts [\*0]. (This can only be true for a SERE accepts\_eword() formula.) Whether the atomic propositions of the formula are all of the form has\_lbt\_atomic\_props() "pnn" where nn is a string of digits. This is required when converting formula into LBT's format.4

### 4.1. Pure Eventualities and Purely Universal Formulas

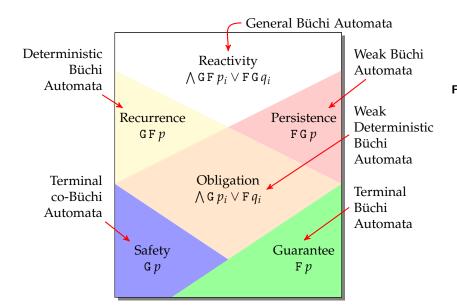
These two syntactic classes of formulas were introduced by? ] to simplify LTL formulas. We shall present the associated simplification rules in Section 5.4.2, for now we only define these two classes.

Pure eventual formulas describe properties that are left-append closed, i.e., any accepted (infinite) sequence can be prefixed by a finite sequence and remain accepted. From an LTL standpoint, if  $\varphi$  is a left-append closed formula, then F  $\varphi \equiv \varphi$ .

Purely universal formulas describe properties that are suffix-closed, i.e., if you remove any finite prefix of an accepted (infinite) sequence, it remains accepted. From an LTL standpoint if  $\varphi$  is a suffix-closed formula, then  $G \varphi \equiv \varphi$ .

<sup>&</sup>lt;sup>3</sup>These "marked" operators are used when translating recurring <>-> or !{r} operators. They are rendered as <>+> and !+{r} and obey the same simplification rules and properties as their unmarked counterpart (except for the is\_marked() property).

<sup>4</sup>http://www.tcs.hut.fi/Software/maria/tools/lbt/



**Figure 4.1.:** The temporal hierarchy of ? ] with their associated classes of automata [? ]. The formulas associated to each class are more than canonical examples: they show the normal forms under which any LTL formula of the class can be rewritten, assuming that p,  $p_i$ , q,  $q_i$  denote subformulas involving only Boolean operators, X, and past temporal operators (Spot does not support the latter).

Let  $\varphi$  denote any arbitrary formula and  $\varphi_E$  (resp.  $\varphi_U$ ) denote any instance of a pure eventuality (resp. a purely universal) formula. We have the following grammar rules:

$$\begin{split} \varphi_E &::= \mathsf{0} \mid \mathsf{1} \mid \mathsf{X} \, \varphi_E \mid \mathsf{X} [!] \, \varphi_E \mid \mathsf{F} \, \varphi \mid \mathsf{G} \, \varphi_E \mid \varphi_E \, \& \, \varphi_E \mid (\varphi_E \mid \varphi_E) \mid ! \, \varphi_U \\ & \mid \varphi \, \mathsf{U} \, \varphi_E \mid \mathsf{1} \, \mathsf{U} \, \varphi \mid \varphi_E \, \mathsf{R} \, \varphi_E \mid \varphi_E \, \mathsf{W} \, \varphi_E \mid \varphi_E \, \mathsf{M} \, \varphi_E \mid \varphi \, \mathsf{M} \, \mathsf{1} \\ \varphi_U &::= \mathsf{0} \mid \mathsf{1} \mid \mathsf{X} \, \varphi_U \mid \mathsf{X} [!] \, \varphi_U \mid \mathsf{F} \, \varphi_U \mid \mathsf{G} \, \varphi \mid \varphi_U \, \& \, \varphi_U \mid (\varphi_U \mid \varphi_U) \mid ! \, \varphi_E \\ & \mid \varphi_U \, \mathsf{U} \, \varphi_U \mid \varphi \, \mathsf{R} \, \varphi_U \mid \mathsf{0} \, \mathsf{R} \, \varphi \mid \varphi_U \, \mathsf{W} \, \varphi_U \mid \varphi \, \mathsf{W} \, \mathsf{0} \mid \varphi_U \, \mathsf{M} \, \varphi_U \end{split}$$

### 4.2. Syntactic Hierarchy Classes

The hierarchy of linear temporal properties was introduced by ? ] and is illustrated on Fig. 4.1. In the case of the LTL subset of the hierarchy, a first syntactic characterization of the classes was presented by ? ], but other presentations have been done including negation [? ] and weak until [? ].

The following grammar rules extend the aforementioned work slightly by dealing with PSL operators. These are the rules used by Spot to decide upon construction to which class a formula belongs (see the methods is\_syntactic\_safety(), is\_syntactic\_guarantee(), is\_syntactic\_obligation(), is\_syntactic\_recurrence(), and is\_syntactic\_persistence() listed on page 14).

The symbols  $\varphi_G$ ,  $\varphi_S$ ,  $\varphi_O$ ,  $\varphi_P$ ,  $\varphi_R$  denote any formula belonging respectively to the Guarantee, Safety, Obligation, Persistence, or Recurrence classes. Additionally  $\varphi_B$  denotes a finite LTL formula (the unnamed class at the intersection of Safety and Guarantee formulas, at the **b**ottom of Fig. 4.1). v denotes any variable, r any SERE,  $r_F$  any bounded SERE (no loops), and  $r_I$  any unbounded SERE.

```
\begin{split} \varphi_{B} &::= 0 \mid 1 \mid v \mid ! \; \varphi_{B} \mid \varphi_{B} \& \; \varphi_{B} \mid (\varphi_{B} \mid \varphi_{B}) \mid \varphi_{B} < -> \varphi_{B} \mid \varphi_{B} \; \text{xor} \; \varphi_{B} \mid \varphi_{B} -> \varphi_{B} \mid \mathsf{X} \; \varphi_{B} \\ & \mid \{r_{F}\} \mid ! \{r_{F}\} \\ \\ \varphi_{G} &::= \varphi_{B} \mid ! \; \varphi_{S} \mid \varphi_{G} \& \; \varphi_{G} \mid (\varphi_{G} \mid \varphi_{G}) \mid \varphi_{S} -> \varphi_{G} \mid \mathsf{X} \; \varphi_{G} \mid \mathsf{F} \; \varphi_{G} \mid \varphi_{G} \, \mathsf{U} \; \varphi_{G} \mid \varphi_{G} \, \mathsf{M} \; \varphi_{G} \\ & \mid ! \; \{r\} \mid \{r\} <> -> \varphi_{G} \mid \{r_{F}\} [] -> \varphi_{G} \\ \\ \varphi_{S} &::= \varphi_{B} \mid ! \; \varphi_{G} \mid \varphi_{S} \& \; \varphi_{S} \mid (\varphi_{S} \mid \varphi_{S}) \mid \varphi_{G} -> \varphi_{S} \mid \mathsf{X} \; \varphi_{S} \mid \mathsf{G} \; \varphi_{S} \, \mathsf{R} \; \varphi_{S} \mid \varphi_{S} \, \mathsf{W} \; \varphi_{S} \\ & \mid \{r\} \mid \{r_{F}\} <> -> \varphi_{S} \mid \{r\} [] -> \varphi_{S} \\ \\ \varphi_{O} &::= \varphi_{G} \mid \varphi_{S} \mid ! \; \varphi_{O} \mid \varphi_{O} \, \& \; \varphi_{O} \mid (\varphi_{O} \mid \varphi_{O}) \mid \varphi_{O} < -> \varphi_{O} \mid \varphi_{O} \; \text{xor} \; \varphi_{O} \mid \varphi_{O} -> \varphi_{O} \\ & \mid \mathsf{X} \; \varphi_{O} \mid \varphi_{O} \, \mathsf{U} \; \varphi_{G} \mid \varphi_{O} \, \mathsf{R} \; \varphi_{S} \mid \varphi_{S} \, \mathsf{W} \; \varphi_{O} \mid \varphi_{G} \, \mathsf{M} \; \varphi_{O} \\ & \mid \{r\} \mid ! \; \{r\} \mid \{r_{F}\} <> -> \varphi_{O} \mid \{r_{I}\} <> -> \varphi_{G} \mid \{r_{F}\} [] -> \varphi_{O} \mid \{r_{I}\} [] -> \varphi_{S} \\ \\ \varphi_{P} &::= \varphi_{O} \mid ! \; \varphi_{P} \mid \varphi_{P} \, \& \; \varphi_{P} \mid (\varphi_{P} \mid \varphi_{P}) \mid \varphi_{P} < -> \varphi_{P} \mid \varphi_{P} \, \mathsf{M} \; \varphi_{P} \\ & \mid \{r\} <> -> \varphi_{P} \mid \{r_{F}\} [] -> \varphi_{P} \mid \{r_{I}\} [] -> \varphi_{S} \\ \\ \varphi_{R} &::= \varphi_{O} \mid ! \; \varphi_{P} \mid \varphi_{R} \, \& \; \varphi_{R} \mid (\varphi_{R} \mid \varphi_{R}) \mid \varphi_{R} < -> \varphi_{R} \mid \varphi_{R} \, \mathsf{Xor} \; \varphi_{R} \mid \varphi_{R} -> \varphi_{R} \\ & \mid \mathsf{X} \; \varphi_{R} \mid \mathsf{G} \; \varphi_{R} \, \; \mathsf{U} \; \varphi_{G} \mid \varphi_{R} \, \mathsf{R} \; \varphi_{R} \mid \varphi_{R} \, \; \mathsf{W} \; \varphi_{R} \mid \varphi_{G} \, \mathsf{M} \; \varphi_{R} \\ & \mid \{r\} [] -> \varphi_{R} \mid \{r_{F}\} <> -> \varphi_{R} \mid \{r_{I}\} <> -> \varphi_{G} \\ \end{aligned}
```

It should be noted that a formula can belong to a class of the temporal hierarchy even if it does not syntactically appears so. For instance the formula  $(G(q \mid FGp) \& G(r \mid FG!p)) \mid Gq \mid Gr$  is not syntactically safe, yet it is a safety formula equivalent to  $Gq \mid Gr$ . Such a formula is usually said pathologically safe.

# 5. Rewritings

#### 5.1. Unabbreviations

The 'unabbreviate()' function can apply the following rewriting rules when passed a string denoting the list of rules to apply. For instance passing the string "^ei" will rewrite all occurrences of xor, <-> and ->.

```
"i"
                                f \rightarrow g \equiv (! f) \mid g
"e"
                               f \iff g \equiv (f \& g) \mid ((!g) \& (!f))
                              f \operatorname{xor} g \equiv (f \& ! g) \mid (g \& ! f)
"^" without "e"
                              f \operatorname{xor} g \equiv !(f \leftrightarrow g)
                                     Fe \equiv e
                                                                                when e is a pure eventuality
"F"
                                    Ff \equiv 1Uf
"G"
                                                                                when u is purely universal
                                    Gu \equiv u
"G" without "R"
                                    Gf \equiv ORf
"GR" without "W"
                                    Gf \equiv fWO
"GRW"
                                    Gf \equiv !F!f
"M"
                                  f M e \equiv F(f \& e)
                                                                                when e is a pure eventuality
"M"
                                  f \, \mathbf{M} \, g \equiv g \, \mathbf{U} \, (g \, \mathbf{\&} \, f)
"R"
                                  f R u \equiv u
                                                                                when u is purely universal
"R" without "W"
                                  f R g \equiv g W (f \& g)
"RW"
                                  f R g \equiv g U ((f \& g) \mid G g)
"W"
                                  f W u \equiv G(f \mid u)
                                                                                when u is purely universal
"W" without "R"
                                  f \otimes g \equiv g \otimes (g \mid f)
"WR"
                                  f \otimes g \equiv f \otimes (g \mid G f)
```

Among all the possible rewritings (see Appendix A) the default rules for R, W and M, those were chosen because they are easier to translate in a tableau construction [?, Fig. 11].

Besides the 'unabbreviate()' function, there is also a class 'unabbreviator() that implements the same functionality, but maintains a cache of abbreviated subformulas. This is preferable if you plan to abbreviate many formulas sharing identical subformulas.

### 5.2. LTL simplifier

The LTL rewritings described in the next three sections are all implemented in the 'tl\_simplifier' class defined in spot/tl/simplify.hh. This class implements several caches in order to quickly rewrite formulas that have already been rewritten previously. For this reason, it is suggested that you reuse your instance of 'tl\_simplifier' as much as possible. If you write an algorithm that will simplify LTL formulas, we suggest you accept an optional 'tl\_simplifier' argument, so that you can benefit from an existing instance.

The 'tl\_simplifier' takes an optional 'tl\_simplifier\_options' argument, making it possible to tune the various rewritings that can be performed by this class. These options cannot be changed afterwards (because changing these options would invalidate the results stored in the caches).

### 5.3. Negative normal form

This is implemented by the 'tl\_simplifier::negative\_normal\_form' method.

A formula in negative normal form can only have negation operators (!) in front of atomic properties, and does not use any of the xor, -> and <-> operators. The following rewriting arrange any PSL formula into negative normal form.

Recall that the negated weak closure  $\{r\}$  is actually implemented as a specific operator, so it is not actually prefixed by the  $\{r\}$  operator.

Note that the above rules include the "unabbreviation" of operators "<->", "->", and "xor", correspondings to the rules "ei<sup>î</sup> of function 'unabbreviate() as described in Section 5.1. Therefore it is never necessary to apply these abbreviations before or after 'tl\_simplifier::negative\_normal\_form'.

If the option 'nenoform\_stop\_on\_boolean' is set, the above recursive rewritings are not applied to Boolean subformulas. For instance calling 'tl\_simplifier::negative\_normal\_form' on ! FG( $a \times b$ ) will produce GF(((! a) & (! b)) | ( $a \times b$ )) if 'nenoform\_stop\_on\_boolean' is unset, while it will produce GF(!( $a \times b$ )) if 'nenoform\_stop\_on\_boolean' is set.

### 5.4. Simplifications

The 'tl\_simplifier::simplify' method performs several kinds of simplifications, depending on which 'tl\_simplifier\_options' was set.

The goals in most of these simplification are to:

- remove useless terms and operator.
- move the X operators to the front of the formula (e.g., X G f is better than the equivalent G X f). This is because LTL translators will usually want to rewrite LTL formulas in a kind of disjunctive form:  $\bigvee_i (\beta_i \wedge X \psi_i)$  where  $\beta_i$ s are Boolean formulas and  $\psi_i$ s are LTL formulas. Moving X to the front therefore simplifies the translation.
- move the F operators to the front of the formula (e.g.,  $F(f \mid g)$  is better than the equivalent  $(Ff) \mid (Fg)$ ), but not before X (XFf is better than FXf). Because Ff incurs some indeterminism, it is best to factorize these terms to limit the sources of indeterminism.

Rewritings defined with  $\equiv$  are applied only when tl\_simplifier\_options::favor\_event\_univ' is true: they try to lift subformulas that are both eventual and universal *higher* in the syntax tree. Conversely, rules defined with  $\equiv$  are applied only when favor\_event\_univ' is false: they try to *lower* subformulas that are both eventual and universal.

Currently all these simplifications assume LTL semantics, so they make no differences between X and X[!]. For simplicity, they are only listed with X.

#### 5.4.1. Basic Simplifications

These simplifications are enabled with tl\_simplifier\_options::reduce\_basics'. A couple of them may enlarge the size of the formula: they are denoted using  $\stackrel{\star}{\equiv}$  instead of  $\equiv$ , and they can be disabled by setting the tl\_simplifier\_options::reduce\_size\_strictly' option to true.

#### **Basic Simplifications for Temporal Operators**

The following are simplification rules for unary operators (applied from left to right, as usual). The terms dnf(f) and cnf(f) denote respectively the disjunctive and conjunctive normal forms if f, handling non-Boolean sub-formulas as if they were atomic propositions.

$$\begin{array}{lll} \operatorname{XFG} f \equiv \operatorname{FG} f & \operatorname{F}(f \operatorname{U} g) \equiv \operatorname{F} g & \operatorname{G}(f \operatorname{R} g) \equiv \operatorname{G} g \\ \operatorname{XGF} f \equiv \operatorname{GF} f & \operatorname{F}(f \operatorname{M} g) \equiv \operatorname{F}(f \operatorname{\&} g) & \operatorname{G}(f \operatorname{W} g) \equiv \operatorname{G}(f \operatorname{I} g) \\ \operatorname{FX} f \equiv \operatorname{XF} f & \operatorname{FG}(f \operatorname{\&} \operatorname{X} g) \equiv \operatorname{FG}(f \operatorname{\&} g) & \operatorname{GF}(f \operatorname{I} \operatorname{X} g) \equiv \operatorname{GF}(f \operatorname{I} g) \\ \operatorname{GX} f \equiv \operatorname{XG} f & \operatorname{FG}(f \operatorname{\&} \operatorname{G} g) \equiv \operatorname{FG}(f \operatorname{\&} g) & \operatorname{GF}(f \operatorname{IF} g) \equiv \operatorname{GF}(f \operatorname{I} g) \\ \operatorname{XO} \equiv \operatorname{O} & \operatorname{FG}(f \operatorname{IG} g) \equiv \operatorname{FG}(f \operatorname{IG} g) & \operatorname{GF}(f \operatorname{\&} \operatorname{F} g) \equiv \operatorname{GF}(f \operatorname{\&} \operatorname{F} g) \\ \operatorname{GF} f \stackrel{\star}{\equiv} \operatorname{GF}(\operatorname{dnf}(f)) & \operatorname{FG}(f \operatorname{\&} \operatorname{F} g) \equiv \operatorname{FG} f \operatorname{\&} \operatorname{GF} g & \operatorname{GF}(f \operatorname{IG} g) \equiv \operatorname{GF} f \operatorname{IF} \operatorname{G} g \\ \\ \operatorname{FG} f \stackrel{\star}{\equiv} \operatorname{FG}(\operatorname{cnf}(f)) & \operatorname{FG}(f \operatorname{IF} g) \equiv \operatorname{FG} f \operatorname{IGF} g & \operatorname{GF}(f \operatorname{IG} g) \equiv \operatorname{GF} f \operatorname{IF} \operatorname{G} g \end{array}$$

$$G(f_1 \mid \ldots \mid f_n \mid GF(g_1) \mid \ldots \mid GF(g_m)) \equiv G(f_1 \mid \ldots \mid f_n) \mid GF(g_1 \mid \ldots \mid g_m)$$

Here are the basic rewriting rules for binary operators (excluding  $\mid$  and & which are considered in Spot as n-ary operators). b denotes any Boolean formula.

$$\begin{array}{c} 1 \, \mathrm{U}\, f \equiv \mathrm{F}\, f \\ f \, \mathrm{M}\, 1 \equiv \mathrm{F}\, f \\ (\mathrm{X}\, f) \, \mathrm{U}\, (\mathrm{X}\, g) \equiv \mathrm{X}(f \, \mathrm{U}\, g) \\ (\mathrm{X}\, f) \, \mathrm{M}\, (\mathrm{X}\, g) \equiv \mathrm{X}(f \, \mathrm{M}\, g) \\ (\mathrm{X}\, f) \, \mathrm{M}\, (\mathrm{X}\, g) \equiv \mathrm{X}(f \, \mathrm{M}\, g) \\ (\mathrm{X}\, f) \, \mathrm{U}\, b \stackrel{\star}{\equiv} b \, \mathrm{I}\, \mathrm{X}(b \, \mathrm{M}\, f) \\ (\mathrm{X}\, f) \, \mathrm{U}\, b \stackrel{\star}{\equiv} b \, \mathrm{I}\, \mathrm{X}(b \, \mathrm{M}\, f) \\ (\mathrm{X}\, f) \, \mathrm{M}\, b \stackrel{\star}{\equiv} b \, \&\, \mathrm{X}(b \, \mathrm{U}\, f) \\ f \, \mathrm{U}\, (\mathrm{G}\, f) \equiv \mathrm{G}\, f \\ f \, \mathrm{M}\, (\mathrm{F}\, f) \equiv \mathrm{F}\, f \\ f \, \mathrm{U}\, (g \, \mathrm{I}\, \mathrm{G}(f)) \equiv f \, \mathrm{W}\, g \\ f \, \mathrm{M}\, (g \, \&\, \mathrm{F}(f)) \equiv f \, \mathrm{M}\, g \\ f \, \mathrm{U}\, (g \, \&\, f) \equiv g \, \mathrm{M}\, f \\ f \, \mathrm{M}\, (g \, \mathrm{I}\, f) \equiv g \, \mathrm{U}\, f \end{array}$$

Here are the basic rewriting rules for n-ary operators (& and |):

```
(FGf)\&(FGg) \equiv FG(f\&g)
                                                                           (GFf) \mid (GFg) \equiv GF(f \mid g)
          (X f) & (X g) \equiv X(f & g)
                                                                                (X f) \mid (X g) \equiv X(f \mid g)
       (X f) & (FGg) \equiv X(f \& FGg)
                                                                              (X f) \mid (G F g) \equiv X(f \mid G F g)
          (G f) & (G g) \equiv G(f \& g)
                                                                                (F f) \mid (F g) \equiv F(f \mid g)
                                                                     (f_1 U f_2) \mid (f_1 U f_3) \equiv f_1 U (f_2 \mid f_3)
(f_1 \cup f_2) \& (f_3 \cup f_2) \equiv (f_1 \& f_3) \cup f_2
(f_1 \cup f_2) \& (f_3 \cup f_2) \equiv (f_1 \& f_3) \cup f_2
                                                                      (f_1 \cup f_2) \mid (f_1 \cup f_3) \equiv f_1 \cup (f_2 \mid f_3)
(f_1 \ W \ f_2) \ \& \ (f_3 \ W \ f_2) \equiv (f_1 \ \& \ f_3) \ W \ f_2
                                                                      (f_1 \ W \ f_2) \ | \ (f_1 \ W \ f_3) \equiv f_1 \ W \ (f_2 \ | \ f_3)
(f_1 R f_2) \& (f_1 R f_3) \equiv f_1 R (f_2 \& f_3)
                                                                      (f_1 R f_2) | (f_3 R f_2) \equiv (f_1 | f_3) R f_2
(f_1 R f_2) & (f_1 M f_3) \equiv f_1 M (f_2 & f_3)
                                                                      (f_1 R f_2) | (f_3 M f_2) \equiv (f_1 | f_3) R f_3
(f_1 M f_2) & (f_1 M f_3) \equiv f_1 M (f_2 & f_3)
                                                                      (f_1 M f_2) | (f_3 M f_2) \equiv (f_1 | f_3) M f_3
       (Fg) & (fUg) \equiv fUg
                                                                              (G f) \mid (f \cup g) \equiv f \cup g
       (Fg) & (fWg) \equiv fUg
                                                                              (G f) \mid (f \forall g) \equiv f \forall g
       (F f) & (f R g) \equiv f M g
                                                                              (Gg) \mid (fRg) \equiv fRg
       (\mathbf{F} f) \& (f \mathbf{M} g) \equiv f \mathbf{M} g
                                                                             (Gg) \mid (f Mg) \equiv f Rg
      f \& ((X f) W g) \equiv g R f
                                                                             f \mid ((X f) R g) \equiv g W f
      f \& ((X f) U g) \equiv g M f
                                                                             f \mid ((X f) M g) \equiv g U f
  f \& (g \mid X(g R f)) \equiv g R f
                                                                        f \mid (g \& X(g W f)) \equiv g W f
  f \& (g \mid X(g \bowtie f)) \equiv g \bowtie f
                                                                        f \mid (g \& X(g \cup f)) \equiv g \cup f
```

The above rules are applied even if more terms are presents in the operator's arguments. For instance FG(a) & G(b) & FG(c) & X(d) will be rewritten as X(d & FG(a & c)) & G(b).

The following more complicated rules are generalizations of  $f \& X G f \equiv G f$  and  $f \mid X F f \equiv F f$ :

$$\begin{split} f \& \operatorname{X}(\operatorname{G}(f \& g \& \ldots) \& h \& \ldots) &\equiv \operatorname{G}(f) \& \operatorname{X}(\operatorname{G}(g \& \ldots) \& h \& \ldots) \\ f \mid \operatorname{X}(\operatorname{F}(f) \mid h \mid \ldots) &\equiv \operatorname{F}(f) \mid \operatorname{X}(h \mid \ldots) \end{split}$$

The latter rule for  $f \mid X(F(f) \mid h ...)$  is only applied if all F-formulas can be removed from the argument of X with the rewriting. For instance  $a \mid b \mid c \mid X(F(a \mid b) \mid F(c) \mid Gd)$  will be rewritten to  $F(a \mid b \mid c) \mid X Gd$  but  $b \mid c \mid X(F(a \mid b) \mid F(c) \mid Gd)$  will only become  $b \mid c \mid X(F(a \mid b \mid c) \mid Gd)$ .

Finally the following rule is applied only when no other terms are present in the OR arguments:

$$F(f_1) \mid \ldots \mid F(f_n) \mid GF(g) \equiv F(f_1 \mid \ldots \mid f_n \mid GF(g))$$

#### **Basic Simplifications for SERE Operators**

The following rules, mostly taken from ? ] are not complete yet. We only show those that are implemented.

The following simplification rules are used for the n-ary operators &&, &, and |. The patterns are of course commutative. b or  $b_i$  denote any Boolean formula while r or  $r_i$  denote any SERE.

```
b \&\& r[*i..j] \equiv \begin{cases} b \&\& r & \text{if } i \leq 1 \leq j \\ 0 & \text{else} \end{cases}
                                                                                                          b \& r \stackrel{\star}{=} \begin{cases} b \mid \{b : r\} & \text{if } \varepsilon r_i \\ b : r & \text{if } \varepsilon \not r_i \end{cases}
   b && \{r_1:\ldots:r_n\}\equiv b && r_1 && \ldots && r_n
  b \&\& \{r_1; \ldots; r_n\} \equiv \begin{cases} b \&\& r_i & \text{if } \exists! i, \varepsilon \not r_i \\ b \&\& (r_1 \mid \ldots \mid r_n) & \text{if } \forall i, \varepsilon r_i \\ 0 & \text{else} \end{cases}
\{b_1; r_1\} && \{b_2; r_2\} \equiv \{b_1 \&\& b_2\}; \{r_1 \&\& r_2\}
                                                                                   \{r_1; b_1\} \&\& \{r_2; b_2\} \equiv \{r_1 \&\& r_2\}; \{b_1 \&\& b_2\}
\{b_1:r_1\} && \{b_2:r_2\}\equiv\{b_1 && b_2\}:\{r_1 && r_2\}
                                                                                   \{r_1:b_1\} && \{r_2:b_2\} \equiv \{r_1 && r_2\}:\{b_1 && b_2\}
 \{b_1 \ ; \ r_1\} \ \& \ \{b_2 \ ; \ r_2\} \equiv \{b_1 \ \&\& \ b_2\} \ ; \ \{r_1 \ \& \ r_2\}
 \{b_1:r_1\} \& \{b_2:r_2\} \equiv \{b_1 \&\& b_2\} : \{r_1 \& r_2\} \text{ if } \varepsilon r_1 \wedge \varepsilon r_2
                       first_match(b[*i..j]) \equiv b[*i]
                       first_match(r[*i..i]) \equiv first_match(r[*i])
                     first_match(r[:*i..j]) \equiv first_match(r[:*i])
                first_match(r_1; r_2[*i.j]) \equiv first_match(r_1; r_2[*i])
              first_match(r_1; r_2[:*i..j]) \equiv first_match(r_1; r_2[:*i])
                first_match(r_1:r_2[*i..j]) \equiv first_match(r_1:r_2[*max(i,1)])
              first_match(r_1:r_2[:*i..j]) \equiv first_match(r_1:r_2[:*i])
                             first_match(b; r) \equiv b; first_match(r)
                  first_match(b[*i..j]; r) \equiv b[*i]; first_match(b[*0..j-i]; r)
                   first_match(b[*i..j]:r) \equiv b[*i-1]; first_match(b[*1..j-i+1]:r) if i > 1
                           first_match(r_1; r_2) \equiv first_match(r_1) if \varepsilon r_2
    first_match(first_match(r_1); r_2) \equiv first_match(r_1); first_match(r_2)
    first_match(first_match(r_1):r_2) \equiv first_match(r_1):first_match(1:r_2)
```

Starred subformulas are rewritten in Star Normal Form [?] with:

$$r[*] \equiv r^{\circ}[*]$$

where  $r^{\circ}$  is recursively defined as follows:

$$r^{\circ} = r \text{ if } \varepsilon \not r$$

$$[*0]^{\circ} = 0 \qquad (r_1; r_2)^{\circ} = r_1^{\circ} \mid r_2^{\circ} \text{ if } \varepsilon r_1 \text{ and } \varepsilon r_2$$

$$r[*i..j]^{\circ} = r^{\circ} \text{ if } i = 0 \text{ or } \varepsilon r \qquad (r_1 \& r_2)^{\circ} = r_1^{\circ} \mid r_2^{\circ} \text{ if } \varepsilon r_1 \text{ and } \varepsilon r_2$$

$$(r_1 \& r_2)^{\circ} = r_1^{\circ} \mid r_2^{\circ} \text{ if } \varepsilon r_1 \text{ and } \varepsilon r_2$$

$$(r_1 \& \varepsilon r_2)^{\circ} = r_1 \& \varepsilon r_2$$

Note: the original SNF definition [? ] does not include '&' and '&&' operators, and it guarantees that  $\forall r, \varepsilon \ / r^\circ$  because  $r^\circ$  is stripping all the stars and empty words that occur in r. For instance  $\{a[*]; b[*]; \{[*0] | c\}\}^\circ[*] = \{a | b | c\}[*]$ . Our extended definition still respects this property in presence of '&' operators, but unfortunately not when the '&&' operator is used.

We extend the above definition to bounded repetitions with:

$$r[*i..j] \equiv r^{\square}[*0..j]$$
 if  $\varepsilon r[*i..j]$ ,  $\varepsilon \not r^{\square}$ , and  $j > 1$   
 $r[*i..j] \equiv r^{\square}[*1..j]$  if  $\varepsilon r[*i..j]$ ,  $\varepsilon r^{\square}$  and  $j > 1$   
 $r[*i..j] \equiv r$  if  $\varepsilon r$  and  $j = 1$ 

where  $r^{\square}$  is recursively defined as follows:

$$r^{\square} = r \text{ if } \varepsilon \text{ } f$$

$$[*0]^{\square} = 0 \qquad (r_1 ; r_2)^{\square} = r_1 ; r_2$$

$$r[*i..j]^{\square} = r^{\square}[*\max(1,i)..j] \text{ if } i = 0 \text{ or } \varepsilon r \qquad (r_1 \& r_2)^{\square} = r_1^{\square} \mid r_2^{\square} \text{ if } \varepsilon r_1 \text{ and } \varepsilon r_2$$

$$(r_1 \& r_2)^{\square} = r_1^{\square} \mid r_2^{\square} \qquad (r_1 \& x_2)^{\square} = r_1 \& x_2$$

The differences between  $\square$  and  $\circ$  are in the handling of r[\*i.j] and in the handling of  $r_1$ ;  $r_2$ .

### **Basic Simplifications SERE-LTL Binding Operators**

The following rewritings are applied to the operators  $[] \rightarrow and <>->$ . They assume that b, denote a Boolean formula.

As noted at the beginning for section 5.4.1, rewritings denoted with  $\stackrel{\star}{\equiv}$  can be disabled by setting the tl\_simplifier\_options::reduce\_size\_strictly' option to true.

Here are the basic rewritings for the weak closure and its negation:

$$\{r[*]\} \equiv \{r\}$$

$$\{r;1\} \equiv \{r\} \text{ if } \varepsilon \neq$$

$$\{r;1\} \equiv 1 \text{ if } \varepsilon r$$

$$\{r;1\} \equiv 1 \text{ if } \varepsilon r$$

$$\{r;1\} \equiv \{r\} \text{ if } \varepsilon \neq$$

$$\{r;1\} \equiv 1 \text{ if } \varepsilon r$$

$$\{r;1\} \equiv \{r\} \text{ if } \varepsilon \neq$$

$$\{r;1\} \equiv \{r\} \text{ i$$

### 5.4.2. Simplifications for Eventual and Universal Formulas

The class of pure eventuality and purely universal formulas are described in section 4.1.

In the following rewritings, we use the following notation to distinguish the class of subformulas:

$$f$$
,  $f_i$ ,  $g$ ,  $g_i$  any PSL formula  $e$ ,  $e_i$  a pure eventuality  $u$ ,  $u_i$  a purely universal formula  $q$ ,  $q_i$  a pure eventuality that is also purely universal

$$\begin{array}{lll} \mathbf{F}\,e \equiv e & f\,\mathbf{U}\,e \equiv e & e\,\mathbf{M}\,g \equiv e\,\&\,g & u_1\,\mathbf{M}\,u_2 \stackrel{\star}{\equiv} (\mathbf{F}\,u_1)\,\&\,u_2 \\ \mathbf{F}(u) \mid q \equiv \mathbf{F}(u \mid q) & f\,\mathbf{U}\,(g \mid e) \equiv (f\,\mathbf{U}\,g) \mid e & f\,\mathbf{M}\,(g\,\&\,u) \equiv (f\,\mathbf{M}\,g)\,\&\,u & q\,\mathbf{U}\,\mathbf{X}\,f \equiv \mathbf{X}(q\,\mathbf{U}\,f) \\ & f\,\mathbf{U}\,(g\,\&\,q) \equiv (f\,\mathbf{U}\,g)\,\&\,q & (f\,\&\,q)\,\mathbf{M}\,g \equiv (f\,\mathbf{M}\,g)\,\&\,q & \\ \mathbf{G}\,u \equiv u & u\,\mathbf{W}\,g \equiv u \mid g & f\,\mathbf{R}\,u \equiv u & e_1\,\mathbf{W}\,e_2 \stackrel{\star}{\equiv} (\mathbf{G}\,e_1) \mid e_2 \\ \mathbf{G}(e)\,\&\,q \equiv \mathbf{G}(e\,\&\,q) & f\,\mathbf{W}\,(g \mid e) \equiv (f\,\mathbf{W}\,g) \mid e & f\,\mathbf{R}\,(g\,\&\,u) \equiv (f\,\mathbf{R}\,g)\,\&\,u & q\,\mathbf{R}\,\mathbf{X}\,f \equiv \mathbf{X}(q\,\mathbf{R}\,f) \\ \mathbf{X}\,q \equiv q & q\,\&\,\mathbf{X}\,f \equiv \mathbf{X}(q\,\&\,f) & q\,\mid\,\mathbf{X}\,f \equiv \mathbf{X}(q\mid f) & \\ \mathbf{X}\,(q\,\&\,f) \equiv q\,\&\,\mathbf{X}\,f & \mathbf{X}\,(q\mid f) \equiv q\mid\mathbf{X}\,f & \\ \end{array}$$