A Unifying Splitting Framework

Gabriel Ebner
(joint work with Jasmin Blanchette and Sophie Tourret)

Vrije Universiteit Amsterdam

2021-06-09

Splitting

$$\frac{C \vee D}{C \quad D}$$

(if C and D are clauses with disjoint variables)

Splitting approaches

- Splitting without backtracking (Riazanov and Voronkov 2001)
- Labelled splitting (Fietzke and Weidenbach 2009)
- Avatar (Voronkov 2014)
  - Very effective: solves 421 previously unsolved problems

Example

Avatar example in our notation

Saturation framework

- General conditions for completeness of saturation provers

- Bachmair and Ganzinger 2001 (handbook article)
- Waldmann et al. 2020

Formulas and consequences

- Abstract formulas $\mathbf{F} = \{C, D, \ldots\}$
- $\bot \in \mathbf{F}$
- Consequence relation $M \models C \quad (M \subseteq \mathbf{F})$

Formulas and consequences

- Abstract formulas $\mathbf{F} = \{C, D, \ldots\}$
- $\bot \in \mathbf{F}$
- Consequence relation $M \models C \quad (M \subseteq \mathbf{F})$
  - $\{\bot\} \models C$
  - $\{C\} \models C$
  - $M \models C$ implies $M \cup N \models C$
  - $\forall C \in N\ (M \models C)$ and $N \models C$ implies $M \models C$

Formulas and consequences

- Abstract formulas $\mathbf{F} = \{C, D, \ldots\}$
- $\bot \in \mathbf{F}$
- Consequence relation $M \models C \quad (M \subseteq \mathbf{F})$
  - $\{\bot\} \models C$
  - $\{C\} \models C$
  - $M \models C$ implies $M \cup N \models C$
  - $\forall C \in N\ (M \models C)$ and $N \models C$ implies $M \models C$

- For example: **F** = set of first-order clauses
- $\{p(x) \vee \neg q(x), q(c)\} \models \{p(c)\}$

Formula redundancy

- $Red_{\mathrm{F}} : \mathcal{P}(\mathbf{F}) \to \mathcal{P}(\mathbf{F})$
  - $Red_{\mathrm{F}}(M) \subseteq Red_{\mathrm{F}}(M \cup N)$
  - $N \setminus Red_{\mathrm{F}}(N) \models \bot \iff N \models \bot$
  - $Red_{\mathrm{F}}(M \cup Red_{\mathrm{F}}(M)) = Red_{\mathrm{F}}(M)$

Formula redundancy

- $Red_{\mathrm{F}} : \mathcal{P}(\mathbf{F}) \to \mathcal{P}(\mathbf{F})$
  - $Red_{\mathrm{F}}(M) \subseteq Red_{\mathrm{F}}(M \cup N)$
  - $N \setminus Red_{\mathrm{F}}(N) \models \bot \iff N \models \bot$
  - $Red_{\mathrm{F}}(M \cup Red_{\mathrm{F}}(M)) = Red_{\mathrm{F}}(M)$

- Standard redundancy criterion:
  "redundant if entailed by smaller formulas"
  $C \in Red_{\mathrm{F}}(M) \iff \{D \in M \mid D \prec C\} \models C$
  - $p(c) \in Red_{\mathrm{F}}(\{p(x)\})$
  - $p(t) \in Red_{\mathrm{F}}(\{t = s, p(s)\})$ (assuming $s \prec t$)
  - $p(y) \notin Red_{\mathrm{F}}(\{p(x)\})$
  - $p(x) \vee \neg p(x) \in Red_{\mathrm{F}}(\emptyset)$

Inference "redundancy"

- Inferences $(C_n, \ldots, C_1, D) \in \mathit{Inf}$
- $\mathit{Red}_{\mathrm{I}} : \mathcal{P}(\mathbf{F}) \to \mathcal{P}(\mathit{Inf})$
  - $\mathit{Red}_{\mathrm{I}}(M) \subseteq \mathit{Red}_{\mathrm{I}}(M \cup N)$
  - $\mathit{Red}_{\mathrm{I}}(M \cup \mathit{Red}_{\mathrm{F}}(M)) = \mathit{Red}_{\mathrm{I}}(M)$
  - $D \in M$ implies $(C_n, \ldots, C_1, D) \in \mathit{Red}_{\mathrm{I}}(M)$
- $\mathit{Inf} \setminus \mathit{Red}_{\mathrm{I}}(\cdot)$ = inferences that must be performed

Dynamic completeness

Theorem

If $(Inf, (Red_{\mathrm{F}}, Red_{\mathrm{I}}))$ *is statically complete, then it is also dynamically complete.*

Extension by labels

A-formulas

- Propositional variables $\mathbf{V}$
- $fml(\cdot) : \mathbf{V} \to \mathbf{F} \cup \mathbf{F}_{\sim}$

- A-formula: $C \leftarrow \{a_1, \ldots, a_n\}$
  - $C \in \mathbf{F}$
  - $a_i \in \mathbf{A} = \mathbf{V} \cup \neg\mathbf{V}$

- Intended interpretation: $fml(a_1) \land \cdots \land fml(a_n) \to C$

Interpretations

- $\mathcal{J} \subseteq \mathbf{A}$ is a propositional interpretation
  if it contains exactly one of v or $\neg v$ for all $v \in \mathbf{V}$
- $C \leftarrow A$ is *enabled* in $\mathcal{J}$ if $A \subseteq \mathcal{J}$
- $\lfloor C \leftarrow A \rfloor = C$
- $\lfloor \mathcal{N} \rfloor \supseteq \mathcal{N}_{\mathcal{J}}$ = all enabled formulas in $\mathcal{N}$
- $\bot \leftarrow A$ is called a *propositional clause*
  - $\bot \leftarrow \{a, \neg b\}$ means $\neg fml(a) \vee fml(b)$!
- $\mathcal{N}_{\bot}$ = all propositional clauses in $\mathcal{N}$

Consequence relation

- Assume consequence relations $\models \subseteq \approx\!\!\!\!|\;$ on $\mathbf{F}$
- $\mathcal{M} \models \mathcal{N}$ if and only if $\mathcal{M}_{\mathcal{J}} \models \lfloor \mathcal{N} \rfloor$ for every $\mathcal{J}$ in which $\mathcal{N}$ is enabled
- $\mathcal{M} \mid\!\approx \mathcal{N}$ if and only if $\mathit{fml}(\mathcal{J}) \cup \mathcal{M}_{\mathcal{J}} \mid\!\approx \lfloor \mathcal{N} \rfloor$ for every $\mathcal{J}$ in which $\mathcal{N}$ is enabled.

Redundancy

- Assume redundancy criterion $(FRed_{\mathrm{F}}, FRed_{\mathrm{I}})$ on **F**
- $C \in FRed_{\mathrm{F}}(\mathcal{N}_{\mathcal{J}})$ for all $\mathcal{J} \supseteq A$; or
- exists $C \leftarrow B \in \mathcal{N}$ such that $B \subset A$.

Inference rules

$$\frac{(C_i \leftarrow A_i)_{i=1}^{n}}{D \leftarrow A_1 \cup \cdots \cup A_n}\ \textsc{Base} \qquad \frac{(\bot \leftarrow A_i)_{i=1}^{n}}{\bot}\ \textsc{Unsat}$$

where $\qquad$ where

$$\frac{C_n \cdots C_1}{D}\ \mathit{FInf} \qquad (\bot \leftarrow A_i)_{i=1}^{n} \text{ is unsat}$$

Admissible inference rules

Theorem

The following are sound inference rules w.r.t. $\approx\!\!\!\!|$, and inferences with = are also simplification rules w.r.t. SRed:

$$\frac{C \leftarrow A}{\bot \leftarrow \{\neg a_1, \ldots, \neg a_n\} \cup A \quad (C_i \leftarrow \{a_i\})_{i=1}^{n}}\ \textsc{Split}$$

$$\frac{(\bot \leftarrow A_i)_{i=1}^{n} \quad C \leftarrow A \cup B}{(\bot \leftarrow A_i)_{i=1}^{n} \quad C \leftarrow B}\ \textsc{Trim} \qquad \frac{fml(a) \leftarrow A}{\bot \leftarrow \{\neg a\} \cup A}\ \textsc{Approx}$$

(where $\{\bot \leftarrow A_i\}_{i=1}^{n} \cup \{\bot \leftarrow A\} \approx\!\!\!\!| \{\bot \leftarrow B\}$) $\qquad \cdots$

Levels of refinement

1. Base calculus ($\mathit{FInf}, \mathit{FRed}$)
2. Splitting calculus ($\mathit{SInf}, \mathit{SRed}$)
3. Model-guided prover MG
4. Locking prover L
5. Given-clause procedure AV

Standard completeness

Assume that ($\mathit{FInf}, \mathit{FRed}$) is statically complete.

Theorem

($\mathit{SInf}, \mathit{SRed}$) is statically complete
(and hence also dynamically complete).

Local completeness

Definition
$(\mathcal{N}_i)_i$ is *locally fair* if either

1. $\bot \in \bigcup_i \mathcal{N}_i$ or
2. exists $\mathcal{J} \models (\mathcal{N}_\infty)_\bot$ such that $\mathit{FInf}((\mathcal{N}_\infty)_\mathcal{J}) \subseteq \bigcup_i \mathit{FRed}_\mathrm{I}((\mathcal{N}_i)_\mathcal{J})$

Theorem
If $(\mathcal{N}_i)_i$ is locally fair and $\mathcal{N}_0 \models \{\bot\}$, then $\bot \in \bigcup_i \mathcal{N}_i$.

Model-guided prover

- Derivations $(\mathcal{J}_0, \mathcal{N}_0) \Longrightarrow_{\mathsf{MG}} (\mathcal{J}_1, \mathcal{N}_1) \Longrightarrow_{\mathsf{MG}} \cdots$
- Transition rules:

DERIVE	$(\mathcal{J}, \mathcal{N} \uplus \mathcal{M}) \Longrightarrow_{\mathsf{MG}} (\mathcal{J}, \mathcal{N} \uplus \mathcal{M}')$	if $\mathcal{M} \subseteq \mathit{SRed}_{\mathrm{F}}(\mathcal{N} \uplus \mathcal{M}')$
SWITCH	$(\mathcal{J}, \mathcal{N}) \Longrightarrow_{\mathsf{MG}} (\mathcal{J}', \mathcal{N})$	if $\mathcal{J}' \models \mathcal{N}_\bot$
UNSAT	$(\mathcal{J}, \mathcal{N}) \Longrightarrow_{\mathsf{MG}} (\mathcal{J}, \mathcal{N} \cup \{\bot\})$	if $\mathcal{N}_\bot \approx \{\bot\}$

Topology on interpretations

- Equip the set of propositional interpretations with the product topology
  - “topology of pointwise convergence”
- Clearly homeomorphic to the Cantor space 2^ω
  - Complete metric space
  - Compact
- Every sequence $(\mathcal{J}_i)_i$ has a convergent subsequence $(\mathcal{J}'_i)_i$
- We call $\lim_{i\to\infty} \mathcal{J}'_i$ a limit point[1]
- Evaluating assertions is continuous

[1] in analogy to other “limits” of clause sets

Local redundancy

- $\mathcal{C} \in \mathit{SRed}_{\mathrm{F}}(\mathcal{M})$ captures “global” redundancy
  - can be removed permanently
  - does not depend on model
- $\lfloor \mathcal{C} \rfloor \in \mathit{FRed}_{\mathrm{F}}(\mathcal{M}_{\mathcal{J}})$ captures “local” redundancy
  - can only be removed *temporarily*
  - relative to current model

Locking prover

- Extra set for A-formulas that are locally redundant depending on some finite subset of the model

LIFT $(\mathcal{J}, \mathcal{N}, \mathcal{L}) \Longrightarrow_{\mathsf{L}} (\mathcal{J}', \mathcal{N}' \cup \lfloor\!\lfloor \mathcal{U} \rfloor\!\rfloor, \mathcal{L} \setminus \mathcal{U})$
if $(\mathcal{J}, \mathcal{N}) \Longrightarrow_{\mathsf{MG}} (\mathcal{J}', \mathcal{N}')$ and $\mathcal{U} = \{(B, C \leftarrow A) \in \mathcal{L} \mid B \not\subseteq \mathcal{J}' \text{ and} A \subseteq \mathcal{J}'\}$

LOCK $(\mathcal{J}, \mathcal{N} \uplus \{C \leftarrow A\}, \mathcal{L}) \Longrightarrow_{\mathsf{L}} (\mathcal{J}, \mathcal{N}, \mathcal{L} \cup \{(B, C \leftarrow A)\})$
if $B \subseteq \mathcal{J}$ and $C \in \mathit{FRed}_{\mathrm{F}}(\mathcal{N}_{\mathcal{J}'})$ for all $\mathcal{J}' \supseteq A \cup B$

Counterexamples

Completeness

Required and allowed inferences

Strongly finitary functions

Conditions

Conclusion

- Completeness of splitting provers depends on subtle details
  - Clause-selection strategies that are complete for nonsplitting provers are not necessarily complete for splitting provers
  - Fairness not only requires a minimum of inferences but also a maximum
- Completeness theorem for an Avatar-like given-clause procedure
  - Requires a very very strong restriction on locking
  - No restriction on the models
- Can we reduce the restriction on locking by requiring “regular” sequences of propositional models?