4.4 Knuth-Bendix Completion (KBC)

Given a set E of equations, the goal of Knuth-Bendix completion is to transform E into an equivalent convergent set R of rewrite rules. If R is finite this yields a decision procedure for E. For ensuring termination the calculus fixes a reduction ordering $\succ$ and constructs R in such a way that $\rightarrow_R \subseteq \succ$, i.e., $l \succ r$ for every $l \rightarrow r \in R$. For ensuring confluence the calculus checks whether all critical pairs are joinable.

The completion procedure itself is presented as a set of abstract rewrite rules working on a pair of equations E and rules R: $(E_0;R_0) \Rightarrow_{\mathrm{KBC}} (E_1;R_1) \Rightarrow_{\mathrm{KBC}} (E_2;R_2) \Rightarrow_{\mathrm{KBC}} \ldots$. The initial state is $(E_0, \emptyset)$ where $E = E_0$ contains the input equations. If $\Rightarrow_{\mathrm{KBC}}$ successfully terminates then E is empty and R is the convergent rewrite system for E_0. For each step $(E; R) \Rightarrow_{\mathrm{KBC}} (E'; R')$ the equational theories of $E \cup R$ and $E' \cup R'$ agree: $\approx_{E\cup R} = \approx_{E'\cup R'}$. By $\mathrm{cp}(R)$ I denote the set of critical pairs between rules in R.

Orient $\quad (E \uplus \{s \mathrel{\dot{\approx}} t\}; R) \;\Rightarrow_{\mathrm{KBC}}\; (E; R \cup \{s \rightarrow t\})$
if $s \succ t$

Delete $\quad (E \uplus \{s \approx s\}; R) \;\Rightarrow_{\mathrm{KBC}}\; (E; R)$

Deduce $\quad (E; R) \;\Rightarrow_{\mathrm{KBC}}\; (E \cup \{s \approx t\}; R)$
if $\langle s, t\rangle \in \mathrm{cp}(R)$

Simplify-Eq $\quad (E \uplus \{s \mathrel{\dot{\approx}} t\}; R) \;\Rightarrow_{\mathrm{KBC}}\; (E \cup \{u \approx t\}; R)$
if $s \rightarrow_R u$

R-Simplify-Rule $\quad (E; R \uplus \{s \rightarrow t\}) \;\Rightarrow_{\mathrm{KBC}}\; (E; R \cup \{s \rightarrow u\})$
if $t \rightarrow_R u$

L-Simplify-Rule $\quad (E; R \uplus \{s \rightarrow t\}) \;\Rightarrow_{\mathrm{KBC}}\; (E \cup \{u \approx t\}; R)$
if $s \rightarrow_R u$ using a rule $l \rightarrow r \in R$ so that $s \sqsupset l$, see below.

Trivial equations cannot be oriented and since they are not needed they can be deleted by the Delete rule. The rule Deduce turns critical pairs between rules in R into additional equations. Note that if $\langle s, t\rangle \in \mathrm{cp}(R)$ then $s\, {}_R{\leftarrow} u \rightarrow_R t$ and hence $R \models s \approx t$. The simplification rules are not needed but serve as reduction rules, removing redundancy from the state. Simplification of the left-hand side may influence orientability and orientation of the result. Therefore, it yields an equation. For technical reasons, the left-hand side of $s \rightarrow t$ may only be simplified using a rule $l \rightarrow r$, if $l \rightarrow r$ cannot be simplified using $s \rightarrow t$, that is, if $s \sqsupset l$, where the *encompassment quasi-ordering* $\mathrel{\underset{\sim}{\sqsupset}}$ is defined by $s \mathrel{\underset{\sim}{\sqsupset}} l$ if $s|_p = l\sigma$ for some p and σ and $\sqsupset = \mathrel{\underset{\sim}{\sqsupset}} \setminus \mathrel{\underset{\sim}{\sqsubset}}$ is the strict part of $\mathrel{\underset{\sim}{\sqsupset}}$.

Lemma 4.4.1. $\sqsupset$ is a well-founded strict partial ordering.

Lemma 4.4.2. If $(E;R) \Rightarrow_{KBC} (E';R')$, then $\approx_{E\cup R} = \approx_{E'\cup R'}$.

Lemma 4.4.3. If $(E;R) \Rightarrow_{KBC} (E';R')$ and $\rightarrow_R \subseteq \succ$, then $\rightarrow_{R'} \subseteq \succ$.

Proposition 4.4.4 (Knuth-Bendix Completion Correctness)**.** If the completion procedure on a set of equations E is run, different things can happen:

1. A state where no more inference rules are applicable is reached and E is not empty. $\Rightarrow$ Failure (try again with another ordering?)
2. A state where E is empty is reached and all critical pairs between the rules in the current R have been checked.
3. The procedure runs forever.

In order to treat these cases simultaneously some definitions are needed:

Definition 4.4.5 (Run)**.** A (finite or infinite) sequence $(E_0;R_0) \Rightarrow_{KBC} (E_1;R_1) \Rightarrow_{KBC} (E_2;R_2) \Rightarrow_{KBC} \ldots$ with $R_0 = \emptyset$ is called a *run* of the completion procedure with input E_0 and $\succ$. For a run, $E_\infty = \bigcup_{i\geq 0} E_i$ and $R_\infty = \bigcup_{i\geq 0} R_i$.

Definition 4.4.6 (Persistent Equations)**.** The sets of *persistent equations of rules* of the run are $E_* = \bigcup_{i\geq 0} \bigcap_{j\geq i} E_j$ and $R_* = \bigcup_{i\geq 0} \bigcap_{j\geq i} R_j$.

Note: If the run is finite and ends with E_n, R_n then $E_* = E_n$ and $R_* = R_n$.

Definition 4.4.7 (Fair Run)**.** A run is called *fair* if $\mathrm{CP}(R_*) \subseteq E_\infty$ (i.e., if every critical pair between persisting rules is computed at some step of the derivation).

Goal: Show: If a run is fair and E_* is empty then R_* is convergent and equivalent to E_0. In particular: If a run is fair and E_* is empty then $\approx_{E_0} = \approx_{E_\infty \cup R_\infty} = \leftrightarrow^*_{E_\infty\cup R_\infty} = \downarrow_{R_*}$.

From now on, $(E_0;R_0) \Rightarrow_{KBC} (E_1;R_1) \Rightarrow_{KBC} (E_2;R_2) \Rightarrow_{KBC} \ldots$ is a fair run and R_0 and E_* are empty.

A *proof* of $s \approx t$ in $E_\infty \cup R_\infty$ is a finite sequence $(s_0, \ldots, s_n)$ so that $s = s_0, t = s_n$ and for all $i \in \{1, \ldots, n\}$ it holds:

1. $s_{i-1} \leftrightarrow_{E_\infty} s_i$ or
2. $s_{i-1} \rightarrow_{R_\infty} s_i$ or
3. $s_{i-1}\ {}_{R_\infty}\!\!\leftarrow s_i$.

The pairs (s_{i-1}, s_i) are called *proof steps*. A proof is called a *rewrite proof* in R_* if there is a $k \in \{0, \ldots, n\}$ so that $s_{i-1} \rightarrow_{R_*} s_i$ for $1 \leq i \leq k$ and $s_{i-1}\ {}_{R_*}\!\!\leftarrow s_i$ for $k+1 \leq i \leq n$.

Idea (Bachmair, Derschowitz, Hsiang): Define a well-founded ordering on proofs so that for every proof that is not a rewrite proof in R_* there is an equivalent smaller proof. Consequence: For every proof there is an equivalent rewrite proof in R_*. A *cost* $c(s_{i-1}, s_i)$ is associated with every proof step as follows:

1. If $s_{i-1} \leftrightarrow_{E_\infty} s_i$ then $c(s_{i-1}, s_i) = (\{s_{i-1}, s_i\}, -, -)$ where the first component is a multiset of terms and $-$ denotes an arbitrary (irrelevant) term.
2. If $s_{i-1} \rightarrow_{R_\infty} s_i$ using $l \rightarrow r$ then $c(s_{i-1}, s_i) = (\{s_{i-1}\}, l, s_i)$.
3. If $s_{i-1} \;{}_{R_\infty}\!\!\leftarrow s_i$ using $l \rightarrow r$ then $c(s_{i-1}, s_i) = (\{s_i\}, l, s_{i-1})$.

Proof steps are compared using the lexicographical combination of the multiset extension of the reduction ordering $\succ$, the encompassment ordering $\sqsupset$ and the reduction ordering $\succ$. The cost $c(P)$ of a proof P is the multiset of the cost of its proof steps. The *proof ordering* $\succ_C$ compares the cost of proofs using the multiset extension of the proof step ordering.

Lemma 4.4.8. $\succ_C$ is well-founded ordering.

Lemma 4.4.9. Let P be a proof in $E_\infty \cup R_\infty$. If P is not a rewrite proof in R_* then there exists an equivalent proof P' in $E_\infty \cup R_\infty$ so that $P \succ_C P'$.

Proof. If P is not a rewrite proof in R_* then it contains

1. a proof step that is in E_∞ or
2. a proof step that is in $R_\infty \backslash R_*$ or
3. a subproof $s_{i-1} \;{}_{R_*}\!\!\leftarrow s_i \rightarrow s_{i+1}$ (peak).

It is shown that in all three cases the proof step or subproof can be replaced by a smaller subproof:
Case 1.: A proof step using an equation $s \doteq t$ is in E_∞. This equation must be deleted during the run.

If $s \doteq t$ is deleted using *Orient*:

$$\ldots s_{i-1} \leftrightarrow_{E_\infty} s_i \ldots \quad\Longrightarrow\quad \ldots s_{i-1} \rightarrow_{R_\infty} s_i \ldots$$

If $s \doteq t$ is deleted using *Delete*:

$$\ldots s_{i-1} \leftrightarrow_{E_\infty} s_{i-1} \ldots \quad\Longrightarrow\quad \ldots s_{i-1} \ldots$$

If $s \doteq t$ is deleted using *Simplify-Eq*:

$$\ldots s_{i-1} \leftrightarrow_{E_\infty} s_i \ldots \quad\Longrightarrow\quad \ldots s_{i-1} \rightarrow_{R_\infty} s' \leftrightarrow_{E_\infty} s_i \ldots$$

Case 2.: A proof step using a rule $s \rightarrow t$ is in $R_\infty \backslash R_*$. This rule must be deleted during the run.

If $s \rightarrow t$ is deleted using *R-Simplify-Rule*:

$$\ldots s_{i-1} \rightarrow_{R_\infty} s_i \ldots \quad\Longrightarrow\quad \ldots s_{i-1} \rightarrow_{R_\infty} s' \;{}_{R_\infty}\!\!\leftarrow s_i \ldots$$

If $s \rightarrow t$ is deleted using *L-Simplify-Rule*:

$$\ldots s_{i-1} \rightarrow_{R_\infty} s_i \ldots \quad\Longrightarrow\quad \ldots s_{i-1} \rightarrow_{R_\infty} s' \leftrightarrow_{E_\infty} s_i \ldots$$

Case 3.: A subproof has the form $s_{i-1} \;{}_{R_*}\!\!\leftarrow s_i \rightarrow_{R_*} s_{i+1}$.

If there is no overlap or a non-critical overlap:

$$\ldots s_{i-1} \;{}_{R_*}\!\!\leftarrow s_i \rightarrow_{R_*} s_{i+1} \ldots \quad \Longrightarrow \quad \ldots s_{i-1} \rightarrow^*_{R_*} s' \;{}^*_{R_*}\!\!\leftarrow s_{i+1} \ldots$$

If there is a critical pair that has been added using *Deduce*:

$$\ldots s_{i-1} \;{}_{R_*}\!\!\leftarrow s_i \rightarrow_{R_*} s_{i+1} \ldots \quad \Longrightarrow \quad \ldots s_{i-1} \leftrightarrow_{E_\infty} s_{i+1} \ldots$$

In all cases, checking that the replacement subproof is smaller than the replaced subproof is routine. □

Theorem 4.4.10 (KBC Soundness)**.** Let $(E_0; R_0) \Rightarrow_{KBC} (E_1; R_1) \Rightarrow_{KBC} (E_2; R_2) \Rightarrow_{KBC} \ldots$ be a fair run and let R_0 and E_* be empty. Then

1. every proof in $E_\infty \cup R_\infty$ is equivalent to a rewrite proof in R_*,
2. R_* is equivalent to E_0 and
3. R_* is convergent.

Proof. 1. By well-founded induction on $\succ_C$ using the previous lemma.

2. Clearly, $\approx_{E_\infty \cup R_\infty} = \approx_{E_0}$. Since $R_* \subseteq R_\infty$ this yields $\approx_{R_*} \subseteq \approx_{E_\infty \cup R_\infty}$. On the other hand, by 1. it holds that $\approx_{E_\infty \cup R_\infty} \subseteq \approx_{R_*}$.
3. Since $\rightarrow_{R_*} \subseteq \succ$, R_* is terminating. By 1. it holds that R_* is confluent.

□

Now using the proof of Theorem 3.15.2 termination of $\Rightarrow_{KBC}$ is undecidable.

Corollary 4.4.11 (KBC Termination)**.** Termination of $\Rightarrow_{KBC}$ is undecidable for some given finite set of equations E.

Proof. Using exactly the construction of Theorem 3.15.2 it remains to be shown that all computed critical pairs can be oriented. Critical pairs corresponding to the search for a PCP solution result in equations $f_R(u(x), v(y)) \approx f_R(u'(x), v'(y))$ or $f_R(u'(x), v'(x)) \approx c$. By chosing an appropriate ordering, all these equations can be oriented. Thus $\Rightarrow_{KBC}$ does not produce any unorientable equations. The rest follows from Theorem 3.15.2. □