

THORN / RCC Planning Session – Memo of Conclusions

1. General Result

THORN successfully proves a large class of RCC composition theorems, including deep chains involving $dc/ec/po/pp/ntpp$, with proofs often exceeding 100k inferences but still completing rapidly.

2. DC Group

The dc -driven problems are structurally easy for the planner. They follow a stable pattern: $dc \rightarrow c \rightarrow split \rightarrow ec/o \rightarrow split \rightarrow overlap \rightarrow decomposition \rightarrow refinement$.

3. EC Group Difficulty

The ec/ec and ec/po problems are fundamentally harder. Naive reuse of dc -style plans fails. The difficulty lies in handling overlap without generating unusable branches (especially $e=$ and $pp-1$).

4. Key Insight

Successful plans avoid broad classification lemmas. Instead, they require context-sensitive refinement lemmas tailored to the specific configuration (e.g. ec/ec or ec/po).

5. Proven Strategy (EC/EC)

A successful pattern was:

$ec \rightarrow c \rightarrow (dc \vee c) \rightarrow (ec \vee o) \rightarrow \text{refine overlap} \rightarrow \text{eliminate } pp/pp-1 \rightarrow \text{map to } tpp/tpp-1 \rightarrow \text{reassemble target}$.

6. Failure Modes

- Global overlap decomposition introduces $e=$ branch (blocks proof)
- Local trichotomy without justification fails
- $pp-1$ elimination is possible but expensive (~40M inferences)

7. New Planning Principle

Introduce lemmas that directly produce target-level relations ($po/pp/ntpp$) instead of passing through $pp/pp-1/e=$.

8. Working Hypothesis

The missing ingredient for EC/PO problems is a lemma of the form:

If $ec(x,y) \ \& \ po(y,z) \ \& \ o(x,z)$ then $po(x,z)$ or $tpp(x,z)$ or $ntpp(x,z)$.

9. Engineering Insight

A timeout control (`thorn.timeout N`) is essential for handling expensive lemmas within plans.

10. Performance Observation

THORN handles $\sim 10^5$ inferences quickly, but certain lemmas explode combinatorially ($\sim 10^7$ – 10^8), requiring careful planning and time control.

11. System Limitation

SBCL stack/heap exhaustion occurs on very large proofs, though results still complete if rerun. Scheme/Shen may mitigate this.

12. Overall Conclusion

The planner is not brute-force: it requires structural insight. The EC family demonstrates that different semantic regimes require different planning strategies.