

On fairness in
distributed
automated deduction

/ Presented at STACS 1993, Würzburg, Germany,
February 1993 */*

M. P. Bonacina

J. Hsiang

Dept. of Computer Science

SUNY AT STONY BROOK

Outline

- Contraction-based theorem proving strategies.
- Distributed deduction: the Clause-Diffusion method.
- Distributed fairness.
- Techniques for fairness of Clause-Diffusion strategies.
- Discussion.

Contraction-based strategies

- Contraction rules: subsumption, demodulation / simplification, normalization, tautology elimination.
- Some of the most successful theorem provers are contraction-based: Otter, RRL, SBR3, Reveal...
- Traditional challenge: prove completeness in the presence of contraction rules.

Contraction-based strategies

A new challenge: parallel contraction-based strategies.

Parallelization is challenging because the data base is not static and not even monotonic.

Backward contraction

- Forward vs. backward contraction.
- Backward contraction is a distinctive feature of contraction-based strategies.
- It entails:
 - highly dynamic data base
 - no read-only data
 - pre-processing does not suffice
 - conflicts
 - "backward contraction
bottle neck"

Parallelism at the search level

Concurrent, asynchronous,

loosely coupled deductive processes

develop their own derivations

by working on separate sets of clauses (no conflicts) and

by exchanging clauses as messages.

Success is reached as soon as one of the processes succeeds.

Distributed environment

• Purely distributed:

- asynchronous, loosely-coupled processors (nodes)
- distributed memory
- communication by message-passing.

• Mixed shared-distributed:

combines message-passing with communication through memory.

Partition the search space

At the clause level :

subdivide the data base of clauses.

For all clauses ψ , assign ψ to a process P_i :

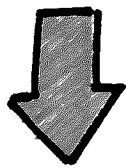
$$S^0 \cup \dots \cup S^i \cup \dots \cup S^{m-1} = S$$

↑
"residents"
at P_i

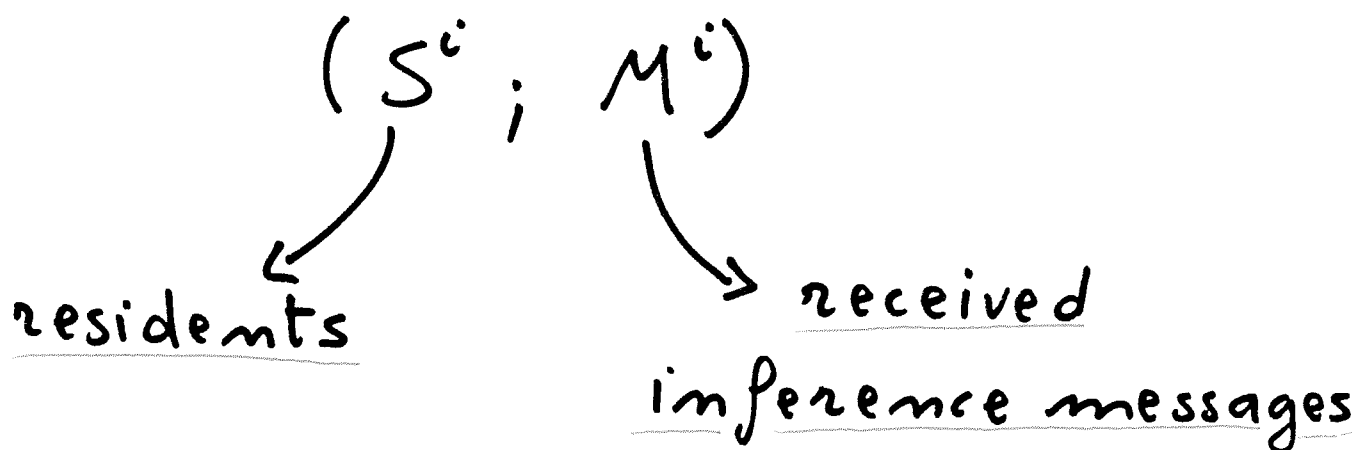
↑
global
data
base

Communication of clauses

Each process p_i takes care of the inferences on S^i , but it is not guaranteed to find a proof by using S^i only.



Each process sends its residents to the others in form of "inference messages"



Partition the search space

At the inference level:

expansion inferences:

if $\psi_1 \in S^i$ and $\psi_2 \in M^i$

P_i parametrizes ψ_2 into ψ_1
but not vice versa.

It also prevents the systematic duplication of expansion steps.

It applies also to other rules,
e.g. resolution, hyper-resolution
and unit-resulting resolution.

No general subdivision of contraction inferences based on ownership of clauses.

In a contraction-based strategy, each process tries to contract as much as possible residents and messages before expansion and communication.

Local contraction (w.r.t. S^i)

and global contraction

(w.r.t. $U S^i$) by schemes for

distributed global contraction.

A Clause - Diffusion strategy

is specified by

- the set of inference rules,
- the search plan which schedules
contraction steps
expansion steps
communication steps
at each process,
- the algorithm to allocate clauses
("new settlers") as residents,
- the mechanism for message-passing,
- the scheme for distributed
global contraction.

Distributed derivation

$(S; M; CP; NS)_0^k \vdash_e$

$(S; M; CP; NS)_1^k \vdash_e \dots$

$\dots \vdash_e (S; M; CP; NS)_i^k \vdash_e \dots$

$1 \leq k \leq n$

S : residents,

M : inference messages,

CP : raw clauses,

NS : new settlers.

Sequential (uniform) fairness

[Huet 81

Bachmain - Ganzinger 92]

$S_0 \vdash_{\epsilon} S_1 \vdash_{\epsilon} \dots \vdash_{\epsilon} S_i \vdash_{\epsilon} \dots$

is uniformly fair if

$$I_{\epsilon}(S_{\infty} - R(S_{\infty})) \subseteq \bigcup_{j \geq 0} S_j.$$

Redundancy criterion R

1) soundness

$$S - R(S) \models R(S).$$

2) monotonicity

$$\text{if } S \subseteq S' \text{ then } R(S) \subseteq R(S').$$

3) if $(S' - S) \subseteq R(S')$

$$\text{then } R(S') \subseteq R(S).$$

[Bachmair, Ganzinger 92]

Distributed (uniform) fairness

$$G_i^k \stackrel{\text{def.}}{=} S_i^k \cup M_i^k \cup CP_i^k \cup NS_i^k$$

$$(S; M; CP; NS)_0^k \quad \tau_e \quad \dots$$

$$\dots \tau_e (S; M; CP; NS)_i^k \tau_e \dots$$

$$1 \leq k \leq m$$

is uniformly fair if

$$I_e(G_\infty - R(G_\infty)) \subseteq \bigcup_{k=1}^m \bigcup_{j \geq 0} G_j^k.$$

Distributed (uniform) fairness

$\forall k, 1 \leq k \leq n,$

$$1) M_{\emptyset}^k = CP_{\emptyset}^k = NS_{\emptyset}^k = \emptyset.$$

$$2) \forall \varphi \in (S_{\emptyset}^k - R(S_{\emptyset}^k))$$

if i is the smallest index s.t.

$$\varphi \in \bigcap_{j>i} S_j^k$$

$$\forall l, \forall \psi \in (S_{\emptyset}^l - R(S_{\emptyset}^l))$$

$$\exists l > i \quad \psi \in M_e^k.$$

$$3) I_e(S_{\emptyset}^k - R(S_{\emptyset}^k)) \subseteq \bigcup_{i>0} CP_i^k.$$

Techniques to satisfy the conditions for fairness

- Inference messages.
- Birth-times of residents and re-emission of inference messages.
- Localized image sets.
- Global image set.
- Wake-up calls.

Deletion of redundant inference messages

"Discard Message":

P_k : sender

P_i : receiver $(i \neq k)$

At P_i :

$\langle \psi_1, P_k, a, x \rangle$ $\langle \psi_2, P_k, a, y \rangle$

$\langle \psi_2, P_k, a, y \rangle$

if

either $\psi_1 > \psi_2$

or $\psi_2 \neq \psi_1$ and $y > x$.

Summary and discussion

- Focus on contraction-based strategies
- Parallelism at the search level:
no synchronization.
- The Clause-Diffusion methodology.
- Distributed derivations.
- Distributed fairness and techniques
to realize it.
- Implementation: the Aquarius
prover.