Combination of distributed search and multi-search in Peers-mcd.d

MARIA PAOLA BONACINA
DEPT. OF COMPUTER SCIENCE
THE UNIVERSITY OF IOWA

Overview of Peers-mcd.d

Logic : equational

Mod AC: yes

Inference rules : paramodulation
simplification
functional subsumption

+

Search plans : several
eager-contraction
best-first search

Strategies : contraction-based

Overview of Peers-mcd.d
(continued)

Parallelization methodology : Modified Clause-Diffusion

Sequential base : EQP 0.9d [Bill McCune 1998]

Language : C + MPI

Ancestors : at CADE 1998
CADE 1997
(fastest Robbins)
CADE 1994
(remotely)

New features

Combination of

distributed search
(parallel processes search
different parts of
search space)

and multi-search
(parallel processes search
whole space by different
search plans)

Recall Modified Clause-Diffusion

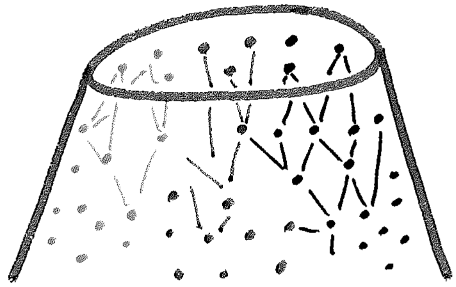
sequential strategy (complete) → [MCD] → distributed strategy (complete)

- Parallel search by N concurrent asynchronous communicating processes.
- Peer processes: no master-slaves.
- N separate derivations: one succeeds ⇒ all halt.
- N separate data bases: separate memories ⇒ no conflicts.

EQP0.9d → [MCD] → Peers-mcd.d

Distributed search in MCD

Subdivide search space:



Dynamic partition of generated clauses ⇒ subdivision of inferences

Every ψ assigned to unique p_i:

$\neg A \vee p = q$ "belongs" to p_i

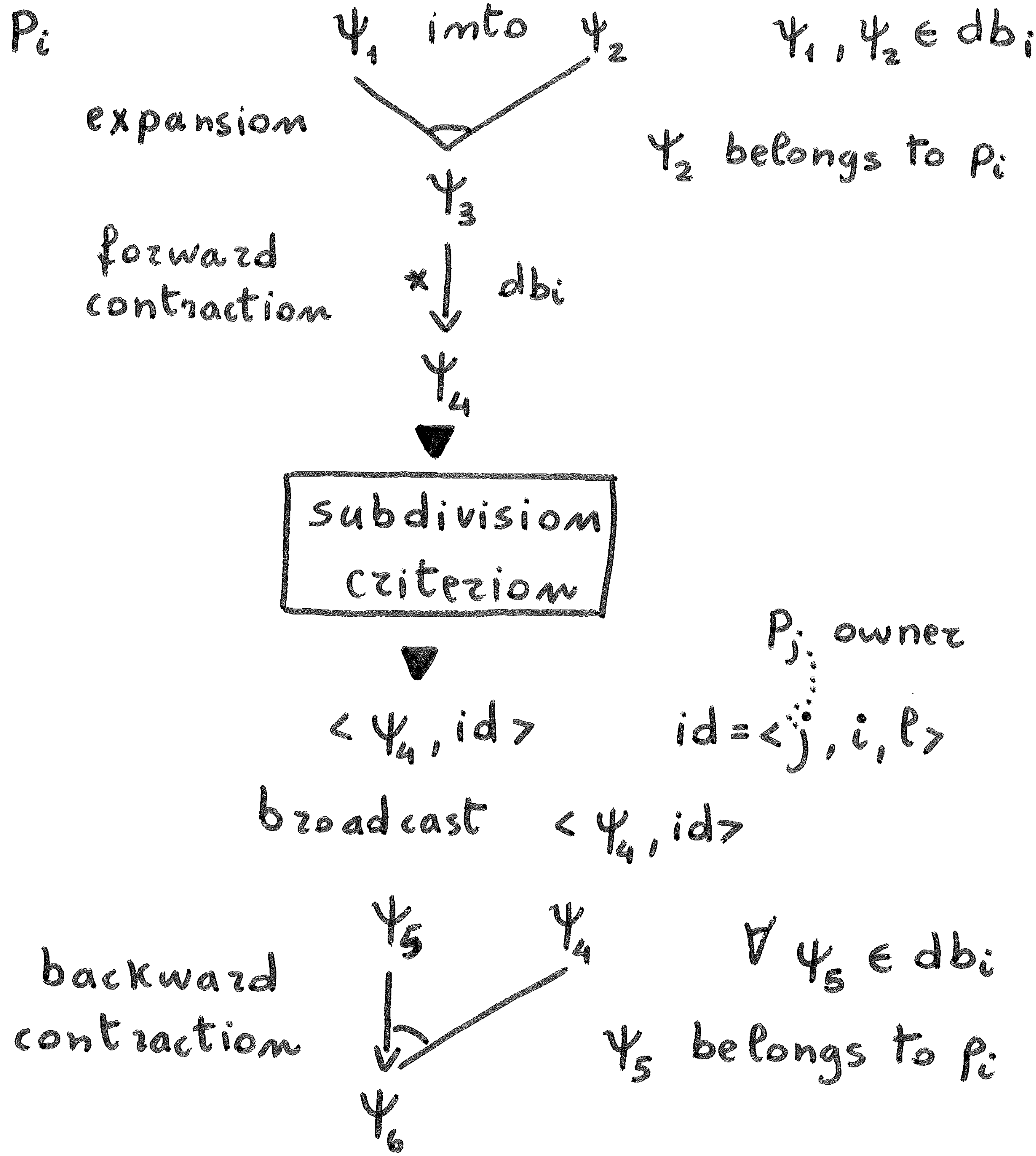
$$\neg A \vee p[s] = q \qquad s = t$$

$$s \not\preceq t$$

$$\neg A \vee p[t] = q$$

allowed only to p_i

Basic mechanism



Remarks

- No need of master / scheduler for subdivision: every process subdivides the clauses it generates.

- No need of master for communication: asynchronous broadcasting.

- Eager forward / backward contraction: Keep each db_i inter-reduced.

- Backward contraction: subdivide simplifications not deletions

Adding multi-search

Three ways to make search plan different:

1) Different premise selection mechanism.

2) Different ratio of breadth-first search and best-first search.

3) Different heuristic function to sort equations for premise selection.

Different premise selection mechanism

Flag DIVERSE-SEL = t

P_i uses
- "given-clause" algorithm if i odd
- "pair" algorithm if i even

Given-clause:
best-first search on equations
heuristic function: weight

ψ — already selected — $\to \varphi_1, \ldots, \to \varphi_n$

Pair: heuristic function: sum of weights
best-first search on pairs of equations $\langle \psi, \varphi \rangle$
$\hookrightarrow \varphi_1 \ldots \varphi_n$

Different breadth-first / best-first ratio

Parameter PICK-GIVEN-RATIO = x : algorithm picks oldest rather than lightest once every $x+1$ choices.

Flag DIVERSE-PICK = 1

P_i resets PICK-GIVEN-RATIO = $x+i$

Different heuristic functions

Flag HEURISTIC-SEARCH = 1

$$P_i \text{ uses } \begin{cases} h_0 & \text{if } i \bmod 3 = 0 \\ h_1 & \text{if } i \bmod 3 = 1 \\ h_2 & \text{if } i \bmod 3 = 2 \end{cases}$$

instead of weight for given-clause.

h_0 = occ-nest
h_1 = cp-in-goal
h_2 = goal-in-cp

} measure syntactic similarity equation / goal

[Siva Anantharaman and Nirina Andrianarivelo 1990]

[Jörg Denzinger and Matthias Fuchs 1994]

Three modes in Peers-mcd.d

- Pure distributed search:
  DECIDE-OWNER-STRAT ≠ NO_SUBDIVIDE
  DIVERSE-SEL ∨ DIVERSE-PICK ∨
  HEURISTIC-SEARCH = 0

- Pure multi-search:
  DECIDE-OWNER-STRAT = NO_SUBDIVIDE
  (broadcast based on heuristic)
  DIVERSE-SEL ∨ DIVERSE-PICK ∨
  HEURISTIC-SEARCH = 1

- Hybrid:
  DECIDE-OWNER-STRAT ≠ NO-SUBDIVIDE
  DIVERSE-SEL ∨ DIVERSE-PICK ∨
  HEURISTIC-SEARCH = 1

Experiments : Moufang identities

Alternative ring (* not associative)

Moufang identities : "approximations" of associativity

First proofs by general-purpose prover:
[Siva Anantharaman & Jieh Hsiang 1990]

using

built-in cancellation laws
[Hsiang - Rusinowitch - Sakai 1987]
inequality ordered saturation rule

⇩

Try without cancellation and IOS.

Left Moufang identity

Mode	*Search plan*	*EQP0.9d*	*1-Peer*	*2-Peers*	*4-Peers*	*6-Peers*	*8-Peers*
D	given(32)	T	T	598	91	187	40
H	given-h(32)	T	415	230	57	42	**9**
D	pair(32)	3,215	3,277	551	109	51	83
D	4-pair(32)	956	1,068	126	**38**	56	58
D	2-pair(32)	88	130	66	39	109	25
H	2d-diverse-h(32)	88	147	84	75	41	25

Average CPU times (in sec)

T = time out after 3600 sec

EQP0.9d on HP B2000/1G

Peers-mcd.d on N HP B2000 or C360 with 1G or 512M

Right Moufang identity

Mode	*Search plan*	*EQP0.9d*	*1-Peer*	*2-Peers*	*4-Peers*	*6-Peers*	*8-Peers*
H	given-h(32)	T	437	268	162	100	**28**
D	pair(32)	T	T	865	356	161	105
H	4d-diverse-h(32)	1,558	1,638	75	**32**	27	47

Super-linear speed-up for all numbers of processes

Best result : 4 processes

speed-up : 48.68

efficiency : 12.17

Middle Moufang identity

EQP 0.9d

"given-clause" algorithm : T

"pair" algorithm : 572

Peers-mcd.d

"given-clause" algorithm :

1-Peer	16
2-Peers	9
4-Peers	5

Discussion of experiments

Pure multi-search: no speed-up.
Heuristic functions improved 1-Peer,
speed-up due to distributed search.

Hybrid mode better than pure
distributed mode:
heuristic functions helped reduce
overlap of processes.

Super-linear speed-up:
much fewer equations generated
(e.g., Right Identity, 4d-diverse-h(32)

EQP0.9d	482,677
2-Peers	122,608)

effective subdivision of space.

Discussion and future work

Study parallelism to provide new forms of search for reasoning.

High-performance deduction needs many tools: parallel search by distributed processes is one.

Design / implementation:
FOL + =
tools for proof comparison,
more experiments.

Theory:

Semantically-guided distributed deduction.