&ACE: A High-Performance Parallel Prolog System for Parallel Symbolic Computing*

Enrico Pontelli[†], **Gopal Gupta**[†], **and Manuel Hermenegildo**[‡]

Abstract: In recent years a lot of research has been invested in parallel processing of numerical applications. However, parallel processing of Symbolic and Artificial Intelligence applications has largely been ignored. This paper presents a high performance system for parallel symbolic computing based on the logic programming paradigm. Logic programming is a paradigm of programming based on a subset of first order logic. An important property of logic programming languages, such as Prolog, is that their declarative semantics is largely independent of the order in which different operations are performed during execution of the program. The different operations can also be performed in parallel, and as a result it is possible to exploit substantial parallelism from Prolog programs, that have been written for sequential machines, without requiring the user to make any modifications to them. In this paper we present a model called ACE that accomplishes this. ACE is a computational model for the full Prolog language, capable of concurrently exploiting both *Or-parallelism* and *Independent And-parallelism*. In this paper we focus on the specific implementation of the and-parallel component of the ACE system (called &ACE) on a shared memory multiprocessor, describing its internal organization, some optimizations to the basic model, and finally presenting some performance figures. These performance figures show that our system is very good at exploiting parallelism from applications written in Prolog since it is able to obtain linear speedups. &ACE is the first ever complete and-parallel system that has been implemented for full Prolog.

Keywords: Parallel Symbolic Computing, Independent And-parallelism, Or-parallelism, Implementation issues, Performance figures

1 Introduction

Parallel processing of numerical problems has been an active area of research for almost two decades now [10, 16, 17]. There are parallelizing compilers and other tools available for parallel numerical processing that programmers can use without being experts on parallel processing. However, the same is not true of parallel processing of Symbolic and Artificial Intelligence (AI) applications. There has been some work [8, 9], but there are still no systems that are general enough and that programmers can use without being experts on parallel processing. In this paper we present a system that is a step towards the goal of building a general automatic tool for parallelizing symbolic and AI applications. Our approach is based on the logic programming paradigm and the Prolog technology. A programmer wishing to parallelize his/her symbolic or AI program will code it in Prolog on a uniprocessor. Our compiler and runtime system will then run this Prolog program automatically in parallel. Our parallelizing system supports the full Prolog language, complete with all its extralogical (the cut), met-

*Some figures in this paper are in color for the convenience of the reviewer. These will be changed to black and white in the final paper, if the paper is accepted.

†Laboratory for Logic, Databases, and Advanced Programming, Dept. of Computer Science, New Mexico State University, Las Cruces, NM, USA.

‡Facultad de Informática, U. Madrid (UPM), Madrid - Spain.

alogical (var, assert, retract, etc.), and side-effect (read, write, etc.) predicates that are sensitive to order of execution of the Prolog program.

Logic programming is a paradigm of programming where programs are expressed as logical implications. Logic programming languages are suited for a wide range of applications, from compilers to databases to symbolic applications, as well as for general purpose programming. The most popular logic programming language is Prolog. An important property of logic programming languages is that they are single assignment languages. Unlike conventional programming languages they disallow destructive assignment and explicit control information. Not only does this allow cleaner (declarative) semantics for programs, and hence a better understanding of them by their users, it also makes it easier for an evaluator of logic programs to employ different control strategies for evaluation. That is, different operations in a logic program can be executed in any order without affecting the (declarative) meaning of the program. In particular, these operations can be performed by the evaluator *in parallel*.

An important characteristic of logic programming languages, thus, is that parallelism can be exploited in an *implicit* way. This can be done directly by the program evaluator (the runtime system) as suggested above, or, alternatively, it can also be done by a parallelizing compiler. The task of the parallelizing compiler is essentially to unburden the evaluator from making run-time decisions regarding which parts of the program to run in parallel. Note that the program can also be parallelized by the user (through suitable annotations). In all cases, the advantage offered by logic programming is that the process is easier because of the more declarative and high level nature of the language. Furthermore, the exploitation of parallelism at run-time or parallelization at compile-time can be done quite successfully in an automatic way and without requiring any input from the user. Clearly, implicit exploitation of parallelism can in many cases have significant advantages over explicit parallelization.[1] In that sense, one can look in some ways towards Prolog for solving a new form of "(parallel) software crisis" that is posed to arise with the new wider availability of multiprocessors[2]—given systems, such as the one described in this paper, one can run Prolog programs written for sequential machines in parallel with no or minimal effort. For the rest of the paper we assume that the reader is familiar with Prolog and its execution model.

It must be pointed out that while the target application area Symbolic and AI applications, in fact, our system can also be used for parallel execution of general purpose (non-numerical) programs. This stems from the fact that Prolog is an excellent language for writing problem solving (non-numerical) programs. This is borne out from some of the benchmarks we have used in the Performance Section (Section 6) of this paper.

Three principal kinds of (implicitly exploitable) control parallelism can be identified in logic programs:

1. *Or-parallelism* arises when more than one clause defines some predicate and a literal unifies with more than one clause head—the corresponding bodies can then be executed in parallel with each other, giving rise to or-parallelism.

2. *Independent and-parallelism* arises when more than one goal is present in the query or in the body of a clause, and the runtime bindings for the variables in these goals are such that two or more goals are *independent* of one another, i.e., their resulting argument terms after applying the bindings of the variables are either variable-free (i.e., *ground*) or have non-intersecting sets of variables. Parallel execution of such goals gives rise to (independent) and-parallelism.

3. *Dependent and-parallelism* arises when two or more goals in the body of a clause have a common variable and are executed in parallel.

In this paper, we are mainly concerned with independent and-parallelism and or-parallelism. The purpose of this paper is to describe in full detail the independent and-parallel component of the ACE system, as it has been developed at the Laboratory for Logic, Databases, and Advanced Programming of the New Mexico State University collaboratively with Universidad Politécnica de Madrid, Spain. In the rest of this paper we describe

[1] This does not mean, of course, that a knowledgeable user should be prevented from parallelizing programs manually or even programming sequentially but in a particular way that makes it possible for the system to uncover more parallelism.

[2] For example, affordable (shared memory) multiprocessor workstations are already being marketed by venders such as Sun (Sun Sparc 10), SGI (Challenge), etc.

the operational semantics of the model of parallelism adopted, details of its implementation, and finally the performance figures obtained on a Sequent Symmetry multiprocessor.

2 ACE: An And-Or Parallel Execution Model

The ACE (And-Or/Parallel Copying-based Execution) model [5] uses stack-copying [1] and recomputation [4] to efficiently support combined Or- and Independent And-parallel execution of logic programs. ACE represents an efficient combination of Or- and independent And-parallelism in the sense that penalties for supporting either form of parallelism are paid only when that form of parallelism is actually exploited. Thus, in the presence of only or-parallelism, execution in ACE is *exactly* as in the MUSE [2] system—a stack-copying based purely Or-parallel system. In the presence of only independent And-parallelism, execution is *exactly* like the &-Prolog [6] system—a recomputation based purely And-parallel system. This efficiency in execution is accomplished by introducing the concept of *teams of processors* and extending the stack-copying techniques of MUSE to deal with this new organization of processors.

2.1 Or-Parallelism in ACE

ACE exploits Or-parallelism by using a stack-copying approach (like MUSE [1, 2]). In this approach, a set of processing *agents* (processors in the case of MUSE, teams of processors in the case of ACE—as explained later) working in or-parallel maintain a *separate* but *identical* address space (i.e. they allocate their data structures starting at the same logical addresses). Whenever an *or-agent* (agents working in or-parallel are termed or-agents) $\mathcal{A}$ is idle, it will start looking for unexplored alternatives generated by some other or-agent $\mathcal{B}$. Once a choice point p with unexplored alternatives is detected in the computation tree $\mathcal{T}_\mathcal{B}$ generated by $\mathcal{B}$, then $\mathcal{A}$ creates a local copy of $\mathcal{T}_\mathcal{B}$ and restarts computation by backtracking over p and executing one of the unexplored alternatives. The fact that all the or-agents maintain an identical logical address space reduces the creation of a local copy of $\mathcal{T}_\mathcal{B}$ to a simple block memory copying operation (Figure 1). This whole operation of obtaining work from another agent is named *sharing* of or-parallel work.

In order to reduce the number of sharing operations performed (since each sharing operation may involve a considerable amount of overhead), unexplored alternatives are always searched starting from the bottommost part of the tree; during the sharing operation all the choice points in between are shared between the two agents (i.e. at each sharing operation we try to maximize the amount of work shared between the two agents). Furthermore, in order to reduce the amount of information transferred during the sharing operation, copying is done *incrementally*, i.e., only the difference between $\mathcal{T}_\mathcal{A}$ and $\mathcal{T}_\mathcal{B}$ is actually copied.

2.2 And-Parallelism in ACE

ACE exploits Independent And-parallelism using a recomputation based scheme[4]—no sharing of solutions is performed (at the and-parallel level). This means that for a query like `?- a,b`, where `a` and `b` are non-deterministic, `b` is completely recomputed for every solution of `a` (as in Prolog). Figure 2 sketches the structure of the computation tree created in the presence of and-parallel computation: a *parbegin-parend* structure is introduced, and the different branches are assigned to different agents. Since we are exploiting only *independent and-parallelism*, only independent subgoals are allowed to be executed concurrently by different *and-agents* (and-agents are processing agents working in and-parallel with each other). Dependencies are detected at run-time by executing some simple tests introduced by the *parallelizing compiler*. In ACE we have adopted the technique originally designed by DeGroot [3] and refined by Hermenegildo [7] (adopted also by &-Prolog [6]) of annotating the program at compile time with *Conditional Graph Expressions (CGEs)*. This will be explained in detail in a later section.

Since and-agents are computing just different parts of the same computation (i.e. they are cooperating in building one solution of the initial query) they need to have *different* but *mutually accessible* logical address spaces.

2.3 And-Or Parallelism in ACE

In ACE a clear separation is made between exploitation of or-parallelism and exploitation of and-parallelism.

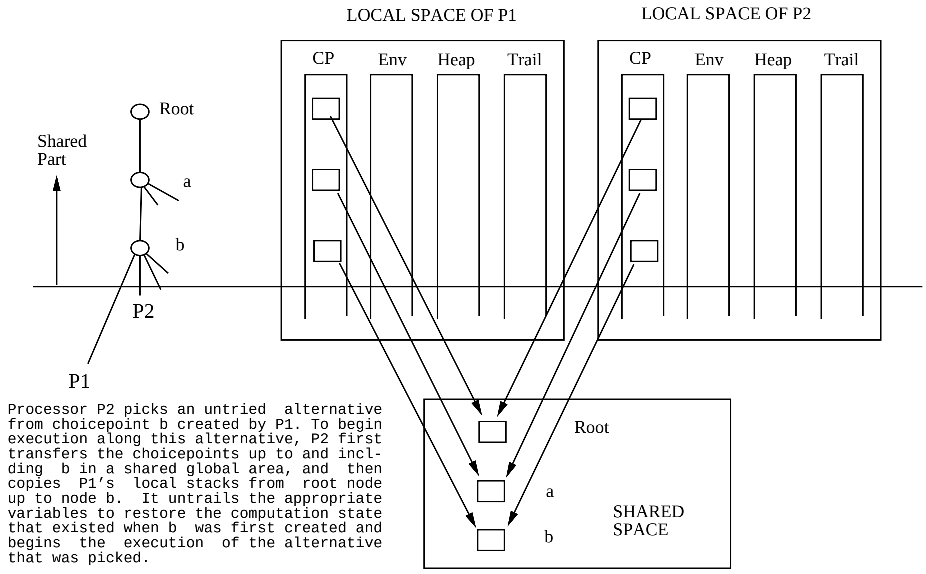

Figure 1: Stack-copying based Or-parallelism

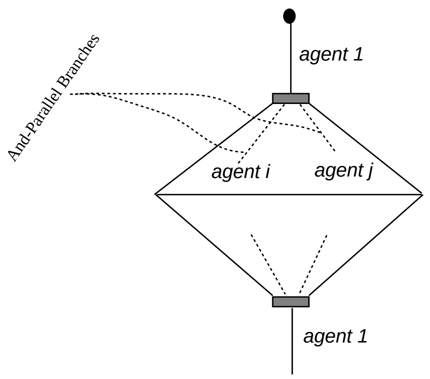

Figure 2: Computation Tree with Recomputation-based And-parallelism

Processors in the multiprocessor system are divided into *teams of processors*. At a higher level, these teams of processors *execute in or-parallel* with each other (i.e., a team will take up only or-parallel work). At a lower level, i.e., within each team, processors in the team *execute in and-parallel* with each other (i.e., along an or-branch taken by a team, processors in that team will execute goals arising in that branch in and-parallel). Thus, the the notion of *or-agent* is mapped to the notion of *team of processors* while the notion of *and-agent* is mapped to the notion of processors inside a team (i.e. each processor is an and-agent). This is illustrated in figure 3.

This organization into teams allows us to: (i) minimize the amount of changes to be done to the basic and-parallel engine; (ii) clearly draw the boundaries between the different components of the system. Each processor in the ACE system is basically an and-parallel engine, capable of carrying on its own computation and interacting with a certain number of other processors (those belonging to the same team). The only new features that need to be added are the followings: (a) a mechanism to keep track of the amount of or-parallel work produced by the computation of a team; (b) a mechanism to allow interaction of one team with another in order to guarantee synchronization and execution of *sharing operations*; (c) extended backtracking, allowing calls to the or-scheduler whenever a team backtracks over a shared choice point.

Or-parallel work is obtained by one team from another through the operation of *sharing*. A sharing operation involves copying (part of) the computation tree generated by a team to another team. The operation is not straightforward since the computation tree is spread over the address spaces of the different processors belonging to the team. Furthermore, we want to perform

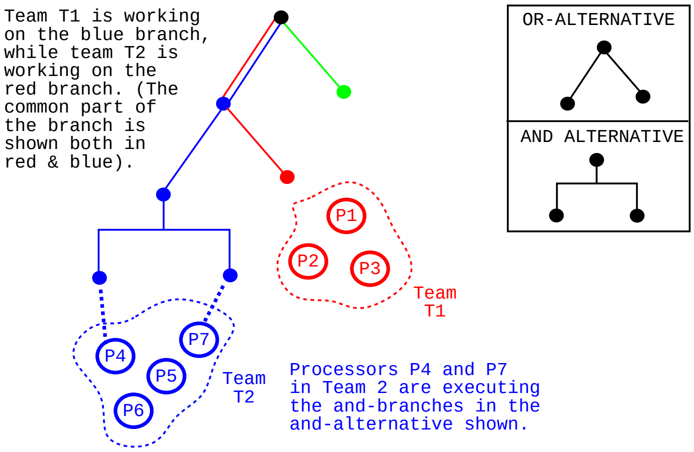

Figure 3: Processor Organization in the ACE system

copying incrementally, i.e., transfer only what is strictly necessary, and this selective operation is complicated by the arbitrary ordering of the different sections of computation on the stacks of the various team members. To make the whole process more effective, we have introduced in ACE some *sharing conditions* that a choice point should satisfy in order to allow its sharing with other teams (i.e., giving away its untried alternatives for or-parallel processing to other teams) . A choice point p satisfies the *sharing conditions* iff the whole computation on its left (in the and-tree exploited by the team) has already been completed. Figure 4 shows an example of this: choice point p satisfies the conditions, while the choice point q does not, since there is a branch on its left which has not completed yet (in the figure, a rectangular box represents an and-parallel call, each branch emanating from such a box corresponds to an and-parallel goal in the and-parallel call; darkened circles represent choicepoints).

In ACE, during a sharing operation, only the choice points satisfying the sharing condition are actually taken into consideration. This simplifies both the copying operation (it is easier to detect which parts of the computation need to be transferred), the scheduling activity (we are guaranteed that everything on the left is terminated and successful and we do not have arbitrary intermixing of shared and private parts in the computation tree), and the management of side-effects and extra-logical predicates.

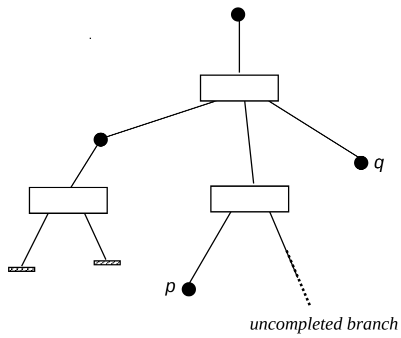

Figure 4: Sharing assumptions in ACE

Different approaches to incremental copying have been studied and heuristics to choose the most appropriate in each situation have been developed. The interested reader is referred to [5] for a detailed discussion of this topic.

3 Independent and-parallelism

The objective of this paper is to illustrate the structure and the features of the and-parallel engine developed for the ACE system. In this section we explain the computational behavior of the and-parallel engine in more detail. It should be pointed out at the outset that our design of the and-parallel component of ACE is heavily influenced by RAP-WAM [7, 6].

3.1 Conditional Graph Expressions

A conditional graph expression (CGE for simplicity) is an expression of the form:

$$(\langle\, \mathit{conditions}\, \rangle \;\Rightarrow\; \mathtt{B}_1 \,\&\, \cdots \,\&\, \mathtt{B}_n\,)$$

where $\langle\, \mathit{conditions}\, \rangle$ is a conjunction of simple tests on variables appearing in the clause (typical tests are *ground*, that verifies whether the argument is instantiated to a ground term, and *independent*, that verifies whether the arguments share any variables with arguments of other goals), and & denotes *parallel conjunction*. The intuitive meaning of a CGE is quite straightforward: if, at run-time, the tests present in *conditions* succeed, then the subgoals $\mathtt{B}_1 \,\&\, \cdots \,\&\, \mathtt{B}_n$ can be executed in and-parallel, otherwise they should be executed sequentially.

A standard Prolog program needs to be annotated with CGEs in order to take advantage of the and-parallel engines available. This process can be done manually by the programmer but is generally done automatically by specialized compile-time analysis tools (like the &-Prolog parallelizing compiler [12], which is also an integral part of ACE).

3.2 Forward Execution

Forward execution of a program annotated with CGEs is quite straightforward. Whenever a CGE is encountered, the conditions are evaluated and, if the evaluation is successful, the various subgoals in the CGE are made available for and-parallel execution. Idle and-agents are allowed to pick up available subgoals and execute them. Only when the execution of those subgoals is terminated the continuation of the CGE (i.e. whatever comes after the CGE) is taken into consideration.

3.3 Backward Execution

Backward execution denotes the series of steps that are performed following a *failure*—due to unification or lack of matching clauses. Since an And-parallel system explores only one or-branch at a time, backward execution involves backtracking and searching for new alternatives in previous choice points. In ACE, where both Or- and And-parallelism are exploited, backtracking should also avoid taking alternatives already taken by other or-agents.

In the presence of CGEs, standard backtracking should be upgraded in order to deal with computations which are spread across processors. &ACE has a complete implementation of such a backtracking scheme.

As long as backtracking occurs over the sequential part of the computation no particular problems occur—we just use plain Prolog-like backtracking. Problem occurs when backtracking involves a CGE. Two cases are possible:

- backtracking occurs **inside** one of the goal $\mathtt{g}_i$ of the CGE, and no solutions are found inside $\mathtt{g}_i$: In this case we can observe that the whole CGE does not have any solution (since the subgoals are known to be independent). This allows the removal of the whole CGE and propagation of backtracking to the computation preceding the CGE itself.
- backtracking occurs in the continuation of the CGE (i.e., **outside** a CGE), and there are no alternatives between the point of failure and the parallel call: In this case backtracking should try to mimic Prolog backtracking, by searching for a new alternative moving from the rightmost goal of the CGE to the leftmost one. If a new successful alternative is found in the goal $\mathtt{g}_i$, then all the subgoals $\mathtt{g}_j$ (with $i < j \leq n$ where n is the number of parallel goals in the CGE) are re-executed in parallel. This is called *recomputation-based* and-parallelism, as mentioned earlier, since the subgoals on the right are completely recomputed for each new solution found on the left. If no successful alternatives are found in any of the subgoals of the CGE, then the whole CGE is removed and backtracking is propagated to the preceding computation.

This scheme permits us to obtain all the solutions of a CGE in the same order in which they are produced in a corresponding sequential execution (of course we are not considering the case in which some of the or-alternatives inside the CGE have been taken by some other or-agent for execution).

4 System Organization

The purpose of this section is to describe in detail the internal organization and the activity of the and-parallel engine of ACE (named &ACE for the sake of simplicity). ACE has been implemented on top of *SICStus Prolog*, one of the most popular implementations of Prolog, and as such it inherits the basic structure of the SICStus WAM architecture [14] together with most of its features and optimizations. The following section is organized as follows: in subsection 4.1 the internal memory layout of &ACE and the data structures allocated during the execution are analyzed. Subsection 4.2 presents a brief overview of the few new instructions added to the SICStus instruction set. Subsection 4.3 deals with the issue of scheduling and-parallel work between the various and-agents. Finally, subsection 4.4 gives an overview of a typical &ACE execution.

4.1 Data Structures

4.1.1 Memory Layout

An and-agent in &ACE keeps a memory organization similar to the one used by a SICStus WAM. Some new areas are added that are described next:

Auxiliary Area: it is used to store various information required during the execution. It is subdivided into two subareas: (i) **Public Area:** this is accessible to all the members of the team, and stores information used to perform and-scheduling (pointers to the top of the goal stack, amount of and-parallel work available, etc.); (ii) **Private Area:** this is accessible only to the specific and-agent and it contains essentially the various registers of the abstract machine (current choice point, current environment, program counter, etc.) and any additional information required by the local execution of the and-agent.

Goal Stack: whenever a CGE is encountered during the execution and the *conditions* are successfully evaluated, parallel execution is activated. The and-agent executing the CGE creates a record (called *goal frame*) for each subgoal of the CGE and stores it in its goal stack. Idle and-agents looking for work will extract a goal frame from the goal stack of a remote agent and use the information stored in it to start the local execution of the subgoal[3].

Descriptors Area: this area, managed as a heap, is used to allocate descriptors of the subgoals executed by the considered and-agent. This is required, as explained in section 5.1.2, to properly manage out-of-sequence kill signals. The other areas are unchanged in their structure w.r.t. SICStus WAM. Note that the architecture of an &ACE agent is very similar to that of an agent in &-Prolog [6].

[3] The structure of a goal frame is illustrated in a later section.

4.1.2 Data Areas

Execution in the WAM is characterized by allocation and deallocation of certain data structures on the stacks of the abstract machine. SICStus WAM uses: (i) the *heap* to allocate *complex terms*; (ii) the *local stack* to allocate *environments*; (iii) the *choice point stack* to allocate *choice points*; (iv) the *trail* to save the addresses relative to *conditional bindings*. &ACE makes use of some additional data structures (also present in RAP-WAM), required to support the and-parallel execution of subgoals. These are as follows:

Parcall Frame: parcall frames are allocated on the choice point stack and are used to identify the execution of a CGE. A parcall frame contains, in addition to other information: (i) a pointer to the environment existing at the time of execution of the CGE (required to allow a proper evaluation of the arguments of each subgoal); (ii) various counters to keep track of the subgoals of the CGE that have been executed, that are still executing, or that are still waiting in the goal stack; (iii) the lock required to guarantee mutual exclusion in the access of common areas (e.g. the parcall frame itself); (iv) a *slot* for each subgoal of the CGE is allocated in the parcall frame, collecting information regarding the execution of the corresponding subgoal (id of the processor which started the subgoal, current status of the execution, id of the processor which completed the execution, etc.).

Goal Frame: goal frames are allocated on the goal stack and are used to maintain the basic information required to start remotely the execution of such subgoal. Typical information stored in a goal frame are: (i) a pointer to the code of the subgoal; (ii) a pointer to the parcall frame relative to the CGE from which the subgoal has been taken. On the goal stack the goal frames are organized in a doubly-linked list. This is required since it may be necessary to remove goal frames that are not lying on the "physical" top of the goal stack (e.g. during a *kill* operation).

Input Marker: input markers are allocated on the choice point stack and they are used to denote the beginning of the execution of a subgoal on a remote processor. An input marker is allocated by the processor that steals the subgoal and it indicates the beginning of a new "section" of the choice point stack (and, thus, of the other areas). Typical information stored in an input marker are: (i) a reference to the parcall frame and to the slot relative to the subgoal; (ii) a pointer to the environment in which the arguments of the subgoal are to be evaluated (i.e. the environment existing at the time in which the parcall frame has been allocated); (iii) a time stamp (used for killing purposes, as explained in section 5.1.2); (iv) a continuation pointer indicating the piece of code to be executed when the subgoal successfully terminates.

End Marker: end markers are allocated on the choice point stack and are used to identify the completion of a subgoal part of a parallel execution. It is used to denote the "logical" end of a stack section and to save the value of some registers which are needed in case of backtracking (like the registers pointing to the current active parcall frame). Since most of the information stored in an end marker is analogous to that required in an input marker (at least the part necessary for register management) and since an input marker can be allocated only after an end-marker, the two structures can be partially merged—in the &ACE implementation the actual input marker allocation phase consists only of filling some empty fields of the previous end marker. This allows a considerable reduction in terms of space and time.

Trail Marker: trail markers are allocated on the choice point stack. To understand the role of the trail markers we need to analyze the two views of the choice points stacks in the &ACE system. Operations on the choice point stack may access the stack following two paths:

1. the *logical* path, in which the data structures relative to the current execution appear to be contiguous on the stack;

2. the *physical* path, in which the data structures relative to the current execution are spread over different parts of the stack, intermixed with (parts of) sections corresponding to the execution of other subgoals.

For example, allocation of a new data structure (e.g. a new choice point) needs to use the physical path (new structures can be allocated only on the real top of the stack), while backtracking should follow the logical path. Intermixing of different computations may occur because of backtracking on trapped subgoals—when the trapped subgoal is reactivated, the new data structures are allocated on the top of the stack—as shown in figure 5: backtracking on the trapped subgoal Y lead to the allocation of a new choice point on the top of the stack; at this point physical and logical paths are different.

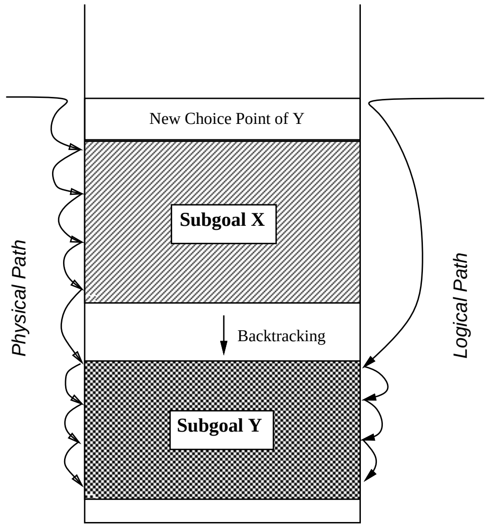

Figure 5: Backtracking on a trapped subgoal

Keeping track of the two paths is quite simple, since each data structure allocated on the choice point stacks contains a pointer to the previous structure along the logical path as well as a a pointer to the previous structure along the physical path[4].

A trail marker is needed in order to keep track of the logical path in some special cases of backtracking. Specifically, a trail marker is required when both of the following conditions are satisfied:

1. we are currently backtracking on a trapped subgoal (i.e. a subgoal which is not lying on the top of the choice point stack);

2. during backtracking on the trapped subgoal we take the last alternative from a choice point and its execution leads to allocation of new data structures.

In this situation a link between the old structures and the new ones is required (mainly for trail management purposes); this link is represented by the trail marker.

[4] Actually this pointer is unnecessary if we know the size of the structure, but keeping it explicitly saves time during the various operations.

As the name suggests, the most relevant pieces of information stored in a trail marker are the current trail pointers.

Section 4.4 illustrates how these data structures are used during an actual execution.

Every structure allocated on the choice point stack includes also some additional information required for trail management and garbage collection.

4.2 Instruction Set

A limited number of changes have been made to the original SICStus WAM instruction set. In general, all the old instructions are unchanged except those needed for additional storing/restoring of the heap pointer. SICStus WAM's `neck` instruction [5] has also been updated to deal with some of the new fields present in the choice point structure.

The new instructions introduced are required to support the execution of the CGEs and the synchronization between parallel executions. These new instructions are as follows:

1. `pcall`: a certain number of instructions have been added to allow creation of a new parallel call (pcall), allocation of a parcall frame, and activation of and-parallel execution.
2. `check_goals`: this instruction typically follows a pcall and is used to support reactivation of parallel execution during outside backtracking;
3. `pop_wait`: this instruction is used to allow the creator of the parallel call to locally execute some of the subgoals of the CGE;
4. `hook`: this instruction represents the "join" of the parallel call. It is executed at the end of a subgoal and allows the and-agent to update the status of the subgoal, allocate the end marker and switch to a different execution.
5. `sch`: this instruction represents the entry-point to the and-scheduler.

The `pcall` and `pop_wait` instructions are inherited from RAP-WAM [7], while the others have been added specifically to support ACE features.

[5] Used in SICStus Prolog to switch from shallow to deep backtracking.

4.3 Scheduling

The and-scheduler is quite simple in its structure. The scheduling algorithm is activated by an idle and-agent and it looks for and-parallel work as follows:

1. if any subgoal is available on the local goal stack, then it will be preferred to any remote work;
2. otherwise other processors are scanned. Priority is given to the processor that signals the highest amount of and-parallel work available.

No restrictions on the subgoal to be selected are imposed. In the future we plan to extend the scheduling algorithm in order to choose work in a "better" way—e.g. selecting subgoals that will permit simpler backtracking mechanisms and a reduced intermixing of executions.

4.4 System Activity

This section briefly illustrates the steps followed by &ACE during a typical execution.

4.4.1 Forward Execution

As long as CGEs are not encountered (i.e. a `pcall` instruction is not executed) forward execution is exactly the same as in the SICStus WAM. When a `pcall` is met, a new parcall frame is allocated, initialized, and all the subgoals but the leftmost one are loaded on the goal stack (the leftmost subgoal is directly executed by the same and-agent that created the parcall frame). Once the lock on the goal stack has been released, all the subgoals become available for parallel execution. At the end of the execution of the leftmost subgoal, the and-agent which started the parcall execution will perform a `pop_wait`, executing locally any further subgoal of the same parcall still available on the local goal stack. Whenever a subgoal is picked up for execution from a goal stack, an input marker is allocated and execution is started. At the end of the execution an end marker is allocated and the information in the parcall frame is appropriately updated. The last and-agent reporting termination of a subgoal of a parcall frame automatically becomes in charge of carrying on the continuation of the CGE. This clearly means that the agent that will complete the execution may be different from the one which started it. An and-agent completing a subgoal while there are still other subgoals running for the same parcall will automatically enter and-scheduling and start searching for new work. Subgoals executed by the various and-agents are uniquely time-stamped (by using a global counter shared among the various and-agents).

4.4.2 Backward Execution

In addition to being able to perform the standard sequential backtracking, backward execution involves:

- being able to perform inside and outside backtracking, as mentioned in section 3.3;
- being able to perform killing of a computation;
- being able to perform communication between and-agents, in order to report situations in which a subgoal should be killed or backtracked over.

Backtracking is performed in the usual way, by moving downwards in the choice point stack and analyzing the data structures encountered. In the case of &ACE this moving downwards is performed following the logical path (not the physical one). The following situations may occur:

1. a *choice point* is encountered: in this case a new alternative is taken and explored. If the new alternative is the last one, the choice point is removed; furthermore if the current subgoal is trapped (i.e. the current point on the logical path does not match the current top of the physical choice point stack), and the new alternative is the last one of the choice point, then a trail marker needs to be allocated to keep track of the new section of trail that we are opening.
2. an *end marker* is encountered: we are entering an outside backtracking situation. If the subgoal below the end marker is the rightmost of the CGE then the end marker is skipped and we keep going down on the same logical path. Otherwise the rightmost subgoal of the CGE is detected and backtracking is transferred to it.
3. an *input marker* is encountered: if the parcall frame referred by the input marker is still in inside status (i.e. we are still looking for the first solution to the CGE) then kill signals are propagated to the other subgoals of the CGE. Immediately after this the agent quits backtracking and moves to the scheduling phase (the agent which created the parcall will carry on the backtracking). If the parcall frame is in outside mode, then reaching an input marker simply means that we have completely explored a subgoal without finding further solutions—backtracking is transferred to the subgoal on its left (using a redo signal to the agent which completed that subgoal).
4. a *parcall frame* is encountered: the situation is exactly as in the previous case (encountering an input marker) since the parcall frame is used as "input marker" for the leftmost subgoal of the CGE.

5 Advanced Issues and Optimizations

5.1 Signal Management

As mentioned in the previous sections, during execution the and-agents need to exchange messages. Each message implies a request sent to the destination agent for execution of a certain activity. The system supports two kinds of messages:

1. *redo* messages—used to request a remote backtracking activity. This is necessary whenever the logical path of the computation continues on the stack of a different agent.
2. *kill* messages—used to request a remote killing activity.

Some of the messages could be avoided by allowing an agent to freely perform backtracking on the stack of another agent We decided to disallow this for the following reasons:

- this would make the integration with Or-parallelism unnecessarily complex;
- the use of explicit messages makes the current status of the execution more "evident", which allows an or-agent looking for work to decide which kind of actions are to be performed in order to appropriately install the stolen alternative;

- desire to keep the overall design simple and clean;
- it is not clear whether this would provide a real advantage, since locking and/or synchronization may become necessary on certain areas of the stack.

Messages are sent and received asynchronously (i.e. a message can be received and served some time after it has been sent but the sender is free to continue its execution immediately after sending the message). The frequency at which an agent checks for the presence of messages can be tuned up by modifying a system defined constant. The examples presented in the last section of this report have been executed allowing the maximum delay between successive message checks (i.e. checks for the presence of messages have been performed only during critical phases of the execution, like opening of new parallel calls, termination of a branch, etc.).

5.1.1 Redo Signals

A redo signal is generated exclusively during outside backtracking and it is used to transfer the current backtracking activity to a subgoal located on the stack of a different agent. To be more precise, the message is sent to the agent which completed the execution of the subgoal on which backtracking has to be performed next. Once received, the agent will locate the end marker relative to that subgoal, remove it and start backtracking from that point (following the logical path relative to that subgoal).

5.1.2 Kill Signals

The management of kill signals is more involved than redo signals. A kill can be generated in two occasions:

1. during inside backtracking a kill message for each subgoal of the CGE to be removed;
2. during outside backtracking a kill message is (or, should be[6]) sent to those subgoals which have shown a deterministic behaviour—since they don't have alternatives there is no sense in trying to search for other solutions.

Details of the *kill* operation are not included due to lack of space and can be found elsewhere [15].

[6] See section 5.2.

5.2 Shallow Parallelism

Innumerable optimizations can be done in a recomputation-based and-parallel system like &ACE. We avoided many of them in order to keep the design of the and-parallel engine simple, a necessary condition for the later introduction of the or-parallel features.

An optimization that has been implemented in the current version of the system deals with taking advantage of deterministic computations. Many of the classical benchmarks proposed for and-parallelism involve the development of deep nestings of parallel calls, while the sequential subgoals (those which do not contain a further parallel call) are deterministic computations. The main idea is that once one of those deterministic computations has been completed, there is no need of keeping any data structure alive (since on backtracking there will not be any alternatives available). For this reason the allocation of the input marker is *delayed* until the first choice point/parcall frame is allocated (in a fashion similar to the shallow backtracking technique); if we reach the end of the computation without allocating any input marker, then the end marker itself is not allocated and we simply record the boundaries of the current trail section in the slot of the parcall frame relative to the current subgoal. On backtracking no kill message needs to be generated for this kind of subgoals (we just need to unwind the trail section indicated in the slot of the parcall frame). This simple optimization allows to save time and space since

- some input and end markers will not be allocated at all on the stack;
- allocation and initialization of data structures on the choice point stack is expensive;
- during backtracking certain kill messages will be avoided.

Figure 6 shows the results of adopting the optimization on the Takeuchi benchmark. Even on a relative small computation time the difference between the unoptimized and the optimized version is significant (on a single processor we have observed an improvement of around 30% for certain benchmarks).

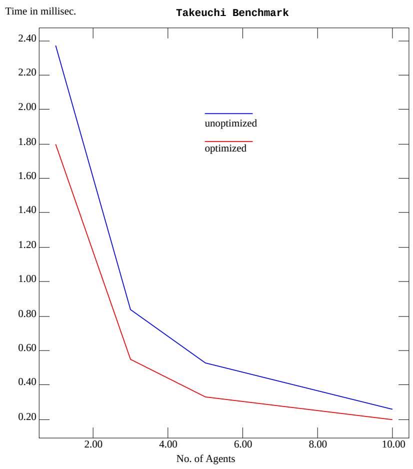

Figure 6: Improvement in execution time using the determinate execution optimization

6 Performance Results

The purpose of this section is to present the results obtained by executing some well-known benchmarks on &ACE. They range from simple test programs to actual applications. The results for the following benchmarks are initially reported: Matrix Multiplication, Quicksort, Takeuchi, Tower of Hanoi, Boyer (a reduced version of the Boyer-Moore theorem prover), Compiler (the Aquarius Prolog compiler from UC Berkeley that is approximately 2,200 lines of Prolog code), Poccur (A list processing program), BT_cluster (A clustering program from British Telecom, UK), Annotator (the annotator part of the &-Prolog parallelizing compiler that is about 1,000 lines) and, Simulator (a simulator for simulating parallel Prolog execution that is about 1,100 lines). Note that our suite of benchmarks does not just contain problems from the area of symbolic computing or AI. Rather, it also contains programs that are very "general" in nature, such as a compiler. This attests to the versatility of Prolog as a language and of the &ACE system that it is able to run in parallel a wide variety of programs.

Table 1 indicates, for each benchmark, the execution time (in ms.) and the speed-up obtained. The execution times are given in the format α/β, where α is the time obtained without the shallow-parallelism optimization and β is the time obtained using the optimized version. The speed-up figures are with respect to the optimized execution.

Table 1 illustrates the speedups obtained for the various benchmarks. The figures clearly indicate that the current implementation, even though not completely optimized, is quite effective. On many benchmarks, containing a sufficient amount of parallelism, the system manages to obtain linear speedups (like matrix multiplication and hanoi). With more processors in the multiprocessor systems we believe we should be able to obtain higher speedups, provided the program contains that much parallel work.

Speedups for some benchmarks are shown in Table 1 and plotted in Figure 7. The largest benchmark is the Aquarius Prolog Compiler (approximately 2,200 lines of Prolog code). Note that for the *compiler*, *quicksort*, and *boyer* benchmarks, the speedup curve flattens out because at some point all available parallelism is exhausted.(e.g., in the case of compiler, the speed-up was measured while the compiler was compiling a program of 8 clauses, thus the maximum speedup that can be obtained will be 8; if a larger program was compiled higher speedups will be obtained). Our implementation incurs an average parallel overhead of about 10% over SICStus Prolog. Some of this parallel overhead is avoided by triggering optimizations mentioned earlier that are based on recognizing determinacy of goals. These optimizations yield, depending on the program, an improvement of 5% to 25% over the unoptimized version. Some improvement data is shown in Table 1 (each entry in Table 1 shows the time in milliseconds before the optimization and after the optimization; the number in parenthesis gives the speed-up obtained; for compiler and simulator benchmarks the unoptimized figure is not shown). As is obvious, improvements due to optimization can be substantial; in some cases superlinear speedup is obtained.

7 Conclusions and Related Work

This paper describes some of the most important features of the independent and-parallel component of the ACE system—a system which implicitly exploits both independent and- and or-parallelism from Prolog programs. We discussed the structure of the machine and the organization of the execution, putting emphasis on new ideas and optimization introduced in the design. We presented results for a large suite of benchmarks, many

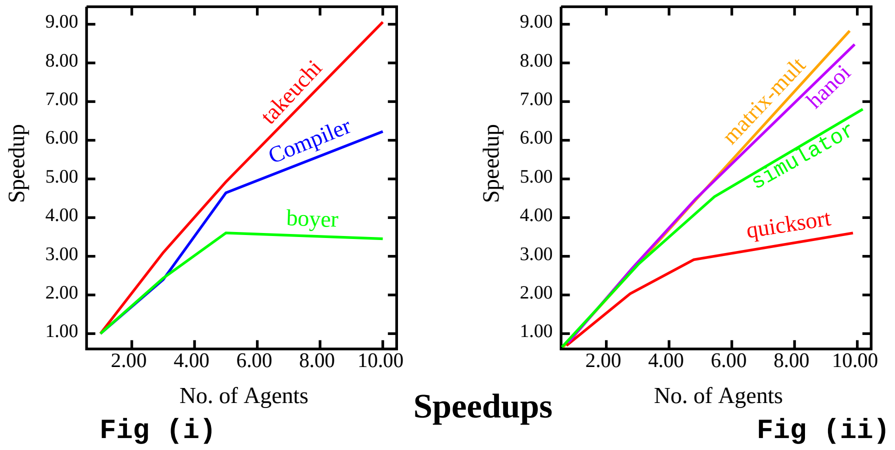

Figure 7: Speedups Curves for Selected Benchmarks on the Sequent Symmetry

Goals executed	&ACE agents 1	3	5	10
matrix_mult(30)	5598/5214	1954/1768 (2.95)	1145/1059 (4.92)	573/534 (9.76)
quick_sort(10)	1882/1536	778/632 (2.43)	548/455 (3.38)	442/373 (4.12)
takeuchi(14)	2366/1811	832/586 (3.09)	521/368 (4.92)	252/200 (9.06)
hanoi(11)	2183/1671	766/550 (3.04)	471/336 (4.97)	231/180 (9.28)
pderiv	—/5375	—/1840 (2.92)	—/1100 (4.89)	—/550 (9.77)
boyer(0)	9655/9290	5329/3829 (2.43)	3816/3199 (2.90)	2887/2687 (3.46)
compiler	—/29902	—/12522 (2.39)	—/6437 (4.65)	—/4801 (6.23)
poccur(5)	3651/3197	1255/1079 (2.96)	759/662 (4.83)	430/371 (8.62)
bt_cluster	1461/1343	528/480 (2.8)	345/312 (4.30)	/189 (7.11)
annotator(5)	1615/1422	556/475 (2.99)	392/322 (4.42)	213/187 (7.60)
simulator	—/3295	—/1232 (2.67)	—/827 (3.98)	—/536 (6.15)

Table 1: Unoptimized/Optimized Execution times in msec (Speedups are shown in parenthesis)

of which are a couple of thousand lines long. Our system is well-suited for parallel execution of symbolic, AI, as well as general purpose applications coded in Prolog. &ACE is a first complete implementation of an and-parallel Prolog engine. The &-Prolog system [6] (to which the &ACE system owes a lot of its inspiration), the system that comes closest to &ACE, however, has only forward execution implemented in it (it should be noted that designing and implementing the parallel backtracking behavior is the most challenging part). As a result not all Prolog programs that can be run in parallel on the &ACE system will run on the &-Prolog system. Lin and Kumar [11] have also implemented an and-parallel system, called APEX, at the University of Texas at Austin. However, they have only tried very simple benchmarks. Their system cannot handle full Prolog[7] and its performance on large benchmarks is not clear. Another and-or parallel system, called Rolog, has been implemented by Ramkumar and Kalé [13] at University of Illinois [13]. There system is based on a non-backtracking model, thus it cannot take advantage of sequential Prolog technology. The sequential execution speed of our system is at least 5 to 10 times better than their's[8]. Rolog also does not

[7] Support for full Prolog is important because not many practical applications can be efficiently written as pure logic programs.

[8] However, Rolog is able to considerably improve its performance by granularity control and parceling out large chunks of computation to efficient Sequential engines such as SICStus 0.6 and Quintus Prolog. Our comparison is based on the assumption

support full Prolog.

8 Acknowledgements

The research presented in this paper has benefited from discussions with Khayri Ali, Manuel Carro, Roland Karlsson, Kish Shen, V. Santos Costa, and David H.D. Warren, all of whom we would like to thank. Ongoing work is supported by NSF Grants CCR 92-11732 and HRD 93-53271, Grant AE-1929 from Sandia National Labs, by an Oak Ridge Associated Universities Faculty Development Award, by NATO Grant CRG 921318, by ESPRIT/CICYT project PARFORCE, and by a fellowship from Phillips Petroleum to Enrico Pontelli.

References

[1] K.A.M. Ali and R. Karlsson. The MUSE Approach to Or-parallel Prolog. *International Journal of Parallel Programming*, 1990. 19(2):129–162.

[2] Khayri A. M. Ali and Roland Karlsson. Full Prolog and Scheduling Or-Parallelism in MUSE. *International Journal of Parallel Programming*, 19(6): 445–475, December 1990.

[3] D. DeGroot. Restricted AND-Parallelism. In *Int'l Conf. on 5th Generation Computer Systems*, pages 471–478. Tokyo, Nov. 1984.

[4] G. Gupta and M. Hermenegildo. Recomputation based Implementation of And-Or Parallel Prolog. In *Int'l Conf. on 5th Generation Computer Sys. '92*, pages 770–782, 1992.

[5] G. Gupta, M. Hermenegildo, E. Pontelli, and V. Santos Costa. ACE: And/Or-parallel Copying-based Execution of Logic Programs. In *Int'l Conf. on Logic Prog.*, MIT Press, pages 93–109, 1994.

[6] M. Hermenegildo and K. Greene. &-Prolog and its Performance: Exploiting Independent And-Parallelism. In *1990 Int'l Conf. on Logic Prog.*, pages 253–268. MIT Press, June 1990.

[7] M. V. Hermenegildo. *An Abstract Machine Based Execution Model for Computer Architecture Design and Efficient Implementation of Logic Programs in Parallel.* PhD thesis, U. of Texas at Austin, August 1986.

[8] L. Kanal and C. B. Suttner (eds). In *Proceedings of the Workshop on Parallel Processing for AI*, 1992. Available from Technische Universität München, Institut für Informatik.

[9] V. Kumar, P.S. Gopalakrishnan, L. N. Kanal (eds). *Parallel Algorithms for Machine Intelligence and Vision.* Springer-Verlag, 1990.

[10] L. Lamport. The parallel execution of DO-loops. In *CACM.* 17, 83-93. 1974.

[11] Y-J. Lin and V. Kumar. AND-parallel execution of Logic Programs on a Shared Memory Multiprocessor: A Summary of Results. In *Proc. Fifth International Logic Programming Conference,* MIT Press. 1988.

[12] K. Muthukumar and M. Hermenegildo. Combined Determination of Sharing and Freeness of Program Variables Through Abstract Interpretation. In *1991 International Conference on Logic Programming.* MIT Press, June 1991.

[13] B. Ramkumar and L. V. Kale. Machine Independent AND and OR Parallel Execution of Logic Programs. Part I and II. In *IEEE Trans. on Par. Distr. Systems* **5**(2):170-192.

[14] Swedish Institute of Computer Science. Industrial Sicstus Prolog Internals Manual, 1989.

[15] D. Tang, E. Pontelli, G. Gupta, and M. Carro. Last Parallel Call Optimization and Fast Backtracking in And-parallel Systems. In *Int'l Logic Prog. Symposium.* MIT Press, to appear, Nov. 1994.

[16] Michael Wolfe. Optimizing Supercompilers for Supercomputers. The MIT Press, 1989.

[17] H. Zima and B. Chapman. Supercompilers for Parallel and Vector Computers. ACM Press. 1991.

that the entire program is run under Rolog.