Proof Development With ΩMEGA: $\sqrt{2}$ Is Irrational

C. Benzmüller

Saarland University, Germany

Linz, Austria, 26. May 2003

Research in the ΩMEGA project

Aim: assistant for the working mathematician

Means: development and integration of heterogenous tools

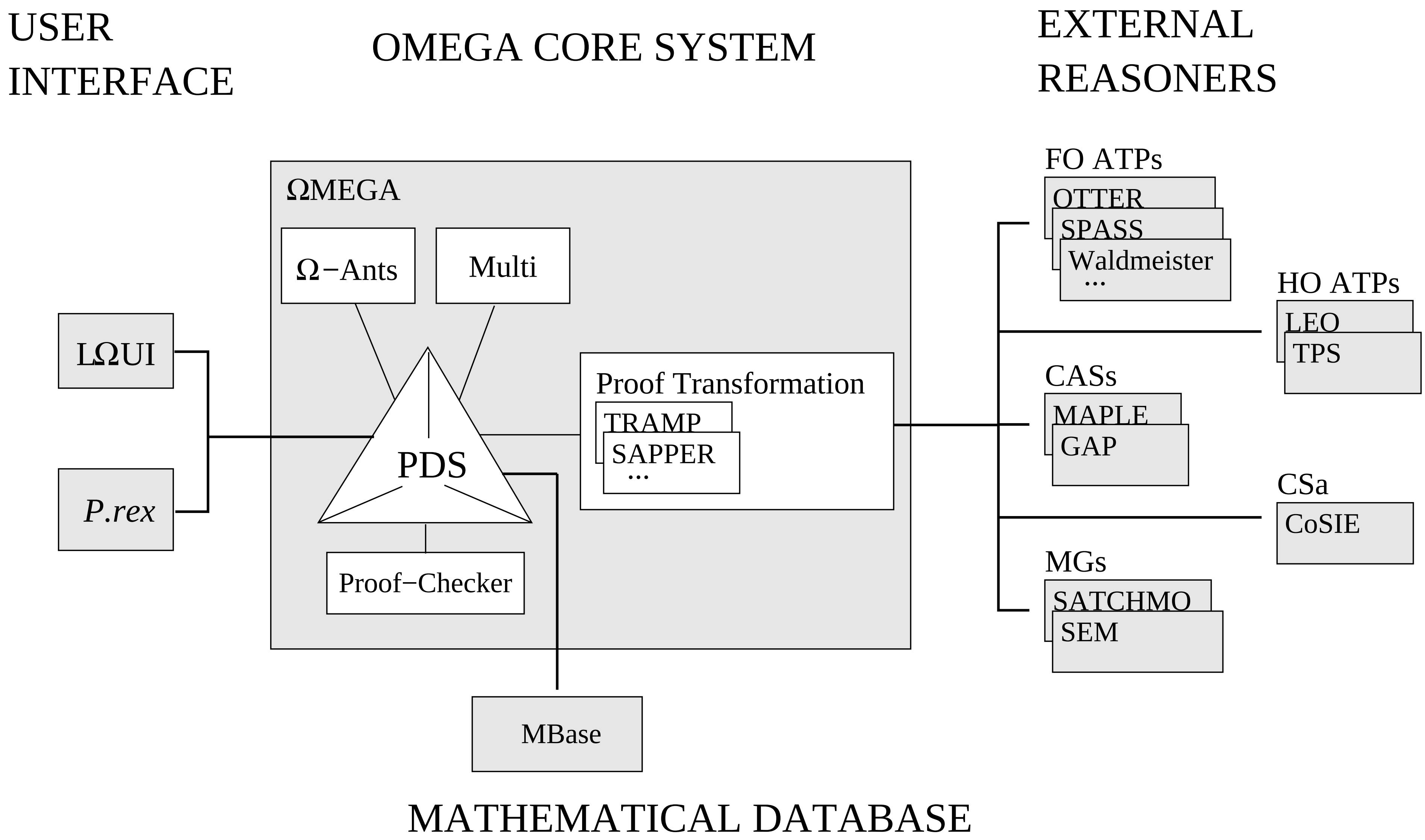
- reasoning — proof planning (PP), agent-based reasoning, ATP
- computation — computer algebra
- interaction — tactical TP, mixed initiative PP
- proof maintenance — proof object, diff. levels of detail
- user interface — graphical UI, natural language
- knowledge management — mathematical database
- infrastructure — network of service systems

ΩMEGA project :=

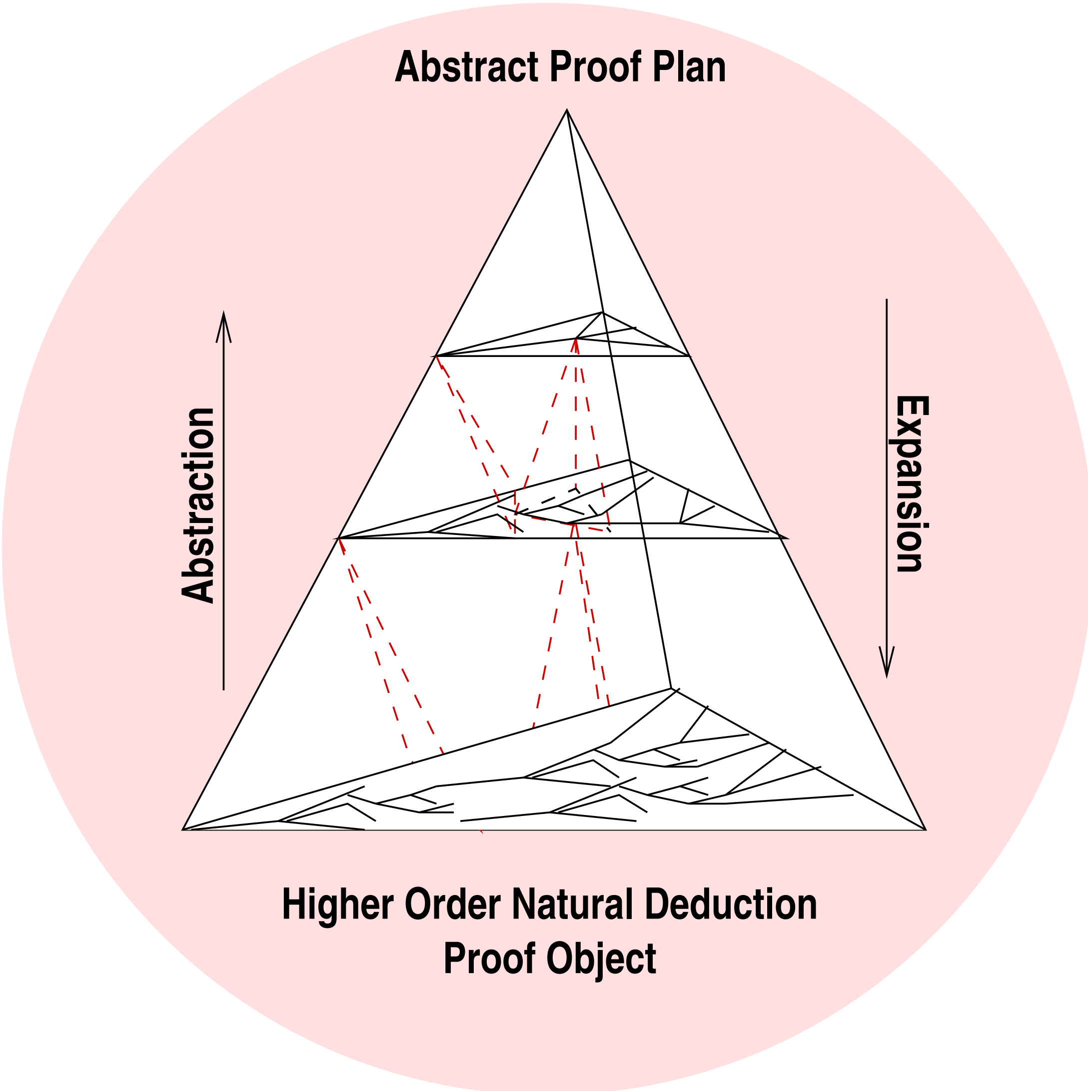
collection of integrated heterogeneous research projects linked

via the core ΩMEGA-system

System Overview



ΩMEGA's Proof Data Structure



Case Study: $\sqrt{2}$ is irrational

- Three contributions:
  1. Tactical theorem proving
  2. New: Interactive island planning
  3. Automated proof planning
- Focus in this talk: Tactical theorem proving and interactive island planning

The $\sqrt{2}$-Problem

Theorem: $\sqrt{2}$ is irrational.

The $\sqrt{2}$-Problem

Theorem: $\sqrt{2}$ is irrational.

Proof: (by contradiction)

Assume $\sqrt{2}$ is rational, that is, there exist natural numbers m, n with no common divisor such that $\sqrt{2} = m/n$. Then $n\sqrt{2} = m$, and thus $2n^2 = m^2$. Hence m^2 is even and, since odd numbers square to odds, m is even; say $m = 2k$. Then $2n^2 = (2k)^2 = 4k^2$, that is, $n^2 = 2k^2$. Thus, n^2 is even too, and so is n. That means that both n and m are even, contradicting the fact that they do not have a common divisor.

The $\sqrt{2}$-Problem

Theorem: $\sqrt{2}$ is irrational.

How closely can we prove the theorem interactively along the previous lines?

Formalization

The Problem:

```
(th~defproblem sqrt2-not-rat (in real)
  (conclusion (not (rat (sqrt 2))))
  (help "sqrt 2 is not a rational number."))
```

Definitions and Lemmas:

```
(th~deftheorem rat-criterion (in real)
  (conclusion
    (forall-sort (lam (x num)
       (exists-sort (lam (y num) (exists-sort (lam (z num)
         (and (= (times x y) z)
              (not (exists-sort (lam (d num)
                         (common-divisor y z d)) int))))
         int)) int)) rat))
  (help "for rationals x there exist integers y,z which

    have no common divisor and z=x*y."))
```

Formalization

```
(th~defdef evenp (in integer)
  (definition
    (lam (x num) (exists-sort (lam (y num) (= x (times 2 y))) int)))
  (help "Definition of even."))

(th~deftheorem square-even (in integer)
  (conclusion
    (forall-sort (lam (x num) (equiv (evenp (power x 2)) (evenp x))) int))
  (help "x is even, iff x^2 is even."))

(th~deftheorem even-common-divisor (in integer)
  (conclusion
    (forall-sort (lam (x num) (forall-sort (lam (y num)
       (implies (and (evenp x) (evenp y)) (common-divisor x y 2)))
     int)) int))
   (help "If x and y are even, then they have a common divisor."))
```

Formalization

```
(th~defdef sqrt (in real)
  (definition
    (lam (x num)
       (choose (lam (y num) (= (power y 2) x)))))
  (help "Definition of square root."))

(th~defdef rat (in rational)
  (definition
    (lam (x num)
        (exists-sort (lam (y num) (exists-sort (lam (z num)
         (and (not (= (mod x y) zero)) (= x (frac y z))))
         pos-nat)) int)))
     (help "Rationals as reduced fractions a/b of integers."))
```

Tactical TP in ΩMEGA

Procedural approach: proof construction by

- applying rules
- applying tactics
  (note difference to LCF style tactics!)
- using external systems
- using facts from the database

Tactical TP in ΩMEGA

Procedural approach: proof construction by

- applying rules
- applying tactics
  (note difference to LCF style tactics!)
- using external systems
- using facts from the database

Verification by proof expansion

Tactical TP in ΩMEGA

Procedural approach: proof construction by

- applying rules
- applying tactics
  (note difference to LCF style tactics!)
- using external systems
- using facts from the database

Verification by proof expansion

Tools for proof presentation

Tactical TP in ΩMEGA

```
OMEGA: load-problems real
;;; Rules loaded for theory REAL.
;;; Theorems loaded for theory REAL.
;;; Tactics loaded for theory REAL.
;;; Methods loaded for theory REAL.
;;; Strategies loaded for theory REAL.
...
OMEGA: prove sqrt2-not-rat
Changing to proof plan SQRT2-NOT-RAT-1
SQRT2-NOT-RAT () |- (NOT (RAT (SQRT 2)))                                   OPEN

OMEGA: noti
NEGATION (NDLINE) A negated line: [SQRT2-NOT-RAT]
FALSITY (NDLINE) A falsity line: [()]

L1 (L1)          |- (RAT (SQRT 2))                                          HYP
L2 (L1)          |- FALSE                                                  OPEN
SQRT2-NOT-RAT () |- (NOT (RAT (SQRT 2)))                              NOTI: (L2)
```

Tactical TP in ΩMEGA

```
OMEGA: import-ass rat-criterion

RAT-CRITERION (RAT-CRITERION) |- (FORALL-SORT ([X].                         THM
                                   (EXISTS-SORT ([Y].
                                     (EXISTS-SORT ([Z].
                                       (AND (= (TIMES X Y) Z)
                                            (NOT (EXISTS-SORT ([D].
                                                   (COMMON-DIVISOR Y Z D))
                                                   INT))))
                                       INT))
                                     INT))
                                   RAT)
```

Tactical TP in ΩMEGA

```
OMEGA: foralle-sort
UNIV-LINE (NDLINE) Universal line: [RAT-CRITERION]
LINE (NDLINE) A line: [()]
TERM (TERM) Term to substitute: (sqrt 2)
SO-LINE (NDLINE) A line with sort: [L1]

L3 (L1) |- (EXISTS-SORT ([DC-248].                          FORALLE-SORT: ((SQRT 2))
             (EXISTS-SORT ([DC-251].                              (RAT-CRITERION L1)
               (AND (= (TIMES (SQRT 2) DC-248) DC-251)
                    (NOT (EXISTS-SORT ([DC-255].
                           (COMMON-DIVISOR DC-248 DC-251 DC-255))
                           INT))))
               INT))
             INT)
```

Tactical TP in ΩMEGA

```
OMEGA: mexistse-sort*
CONCLINE (NDLINE) Conclusion Line.: [L2]
EXLINE (NDLINE) An existentially quantified line: [L3]
SUBGOAL (NDLINE) Subgoal Line.: [()]
PARAMETER (TERMSYM-LIST) Termsym List.: [(dc-248 dc-251)](n m)

L4  (L4)        |- (AND (INT N)                                          HYP
                        (EXISTS-SORT ([DC-251].
                          (AND (= (TIMES (SQRT 2) N) DC-251)
                               (NOT (EXISTS-SORT ([DC-255].
                                      (COMMON-DIVISOR N DC-251 DC-255))
                                      INT))))
                          INT))
L6  (L4)        |- (INT N)                                       ANDEL: (L4)
L5  (L5)        |- (AND (INT M)                                          HYP
                        (AND (= (TIMES (SQRT 2) N) M)
                             (NOT (EXISTS-SORT ([DC-255].
                                    (COMMON-DIVISOR N M DC-255))
                                    INT))))
L8  (L5)        |- ...
```

Tactical TP in ΩMEGA

```
OMEGA: ande
CONJUNCTION (NDLINE) Conjunction to split: [L9]
LCONJ (NDLINE) Left conjunct: [()]
RCONJ (NDLINE) Right conjunct: [()]

L11 (L5) |- (= (TIMES (SQRT 2) N) M)                                  ANDE: (L9)
L12 (L5) |- (NOT (EXISTS-SORT ([DC-255].                              ANDE: (L9)
                   (COMMON-DIVISOR N M DC-255)) INT))
```

Now we are stuck: from L11 we want to infer

`(= (times 2 (power n 2)) (power m 2))`

then `(evenp (power m 2))` and `(evenp m)`

No tactic available for this; instead cut rule is needed

Tactical TP in ΩMEGA

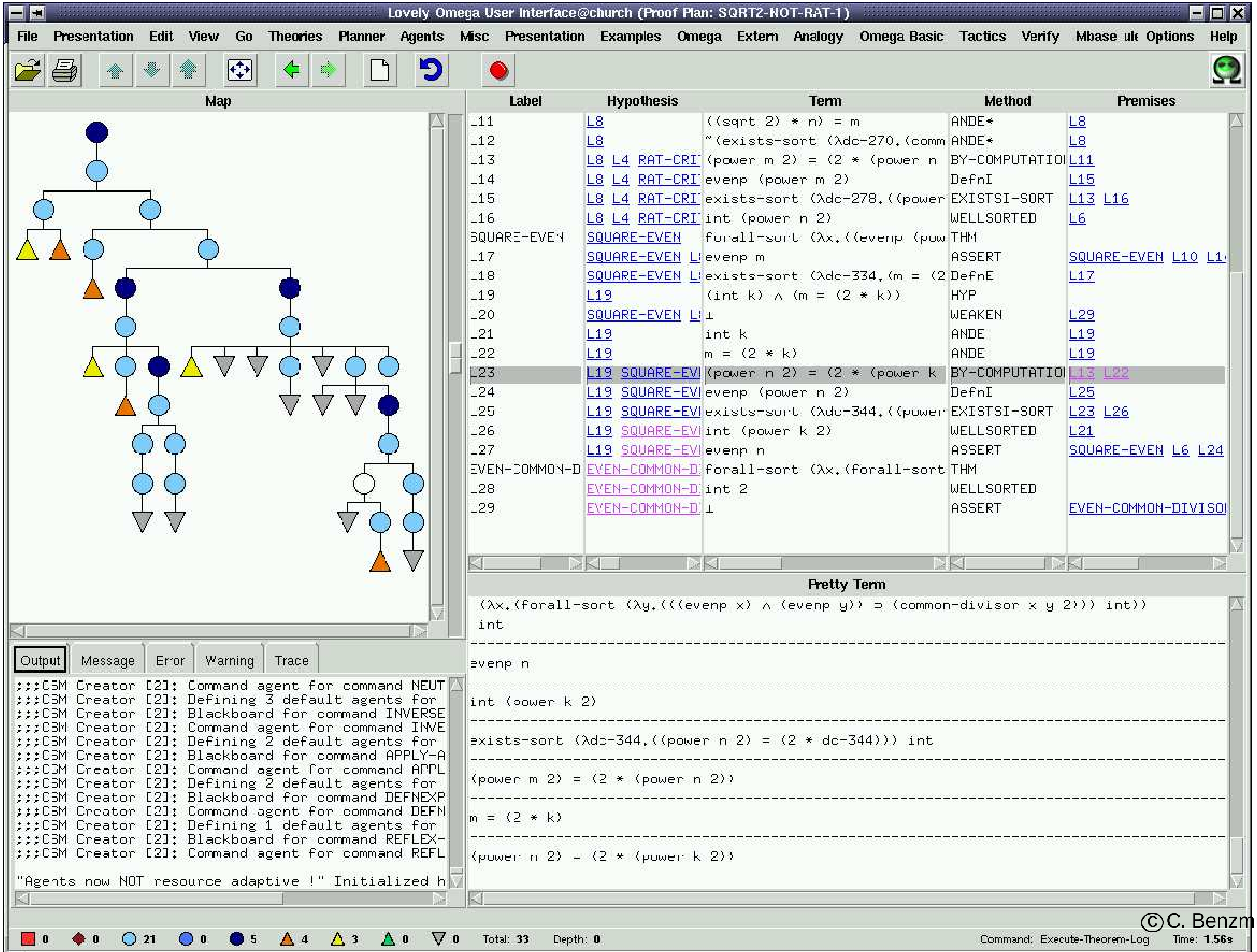
```
OMEGA: lemma
NODE (NDPLANLINE) An open node: [L9]
FORMULA (FORMULA) Formula to be proven as lemma:
(= (power m 2) (times 2 (power n 2)))

L13         (L8 L4 L1) ! (= (POWER M 2) (TIMES 2 (POWER N 2)))      OPEN

OMEGA: by-computation
LINE1 (NDLINE) A line with an arithmetic term to justify.: l13
LINE2 (NDLINE-LIST) A list containing premises to be used.: (l11)

L11         (L8)       ! (= (TIMES (SQRT 2) N) M)             ANDE*: (L8)

L13         (L8 L4 L1) ! (= (POWER M 2)                 BY-COMPUTATION:(L11)
                            (TIMES 2 (POWER N 2)))
```

The latter tactic employs the CAS MAPLE via MATHWEB.

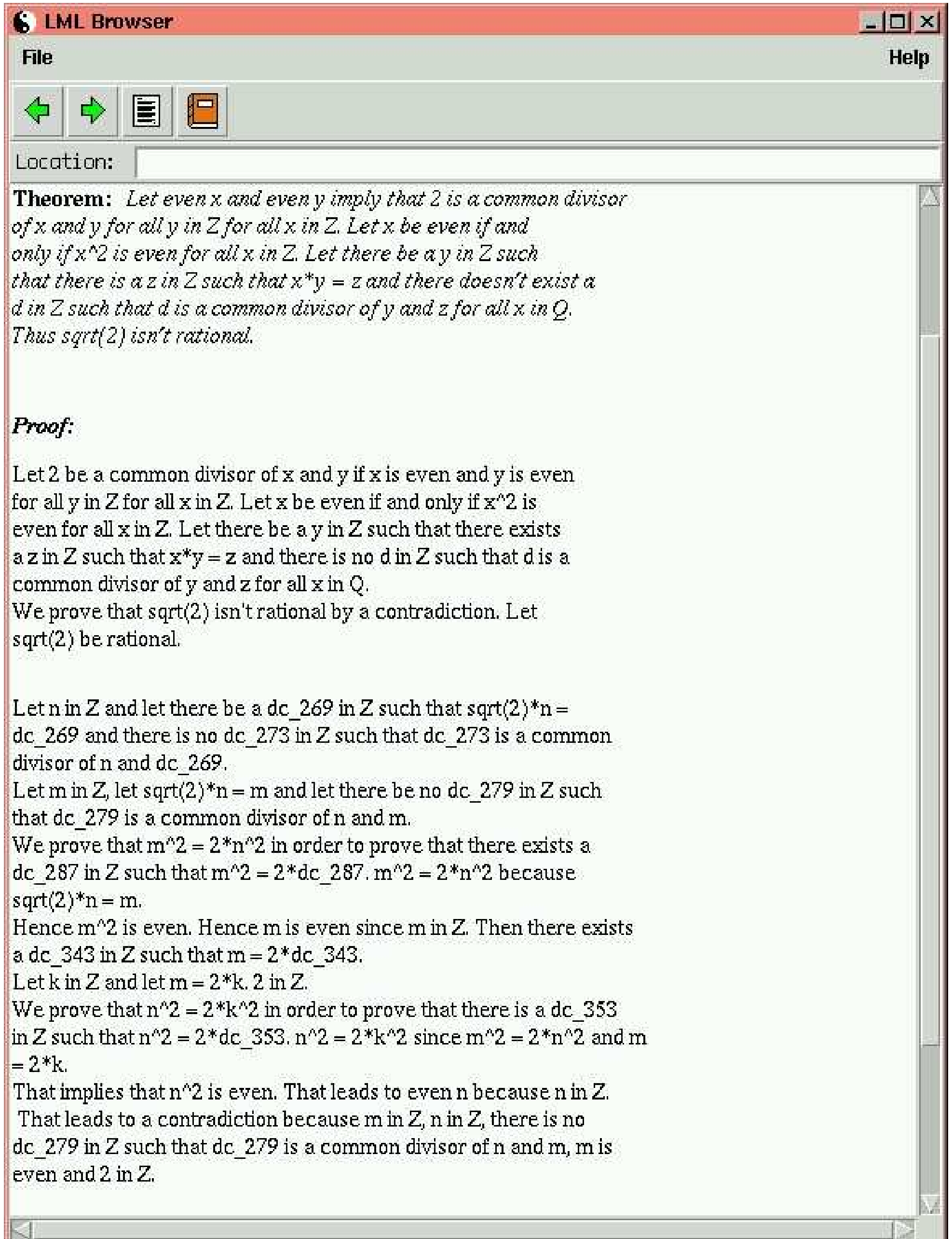
Tactical TP in ΩMEGA

```
...
(L89 (TERTIUM-NON-DATUR POWER-INT-CLOSED NAT-INT SUCC-NAT ZERO-NAT
      EVEN-COMMON-DIVISOR SQUARE-EVEN L4)
     (EQUIV (EVENP (POWER N 2)) (EVENP N))
     (0 ("IMPE" () (L6 L88) "grounded" () ("EXISTENT" "EXISTENT" "EXISTENT")))
      )
(L90 (TERTIUM-NON-DATUR POWER-INT-CLOSED NAT-INT SUCC-NAT ZERO-NAT
      EVEN-COMMON-DIVISOR SQUARE-EVEN L4)
     (IMPLIES (EVENP (POWER N 2)) (EVENP N))
     (2 ("EQUIVE" () (L89) "expanded" () ("EXISTENT" "L91" "EXISTENT"))
        ("ANDE" () (L125) "expanded" () ("EXISTENT" "L91" "EXISTENT"))
        ("ANDEL" () (L125) "grounded" () ("EXISTENT" "EXISTENT")))
      )
...
```

Tactical TP in ΩMEGA



Tactical TP in ΩMEGA



Theorem: *Let even x and even y imply that 2 is a common divisor of x and y for all y in Z for all x in Z. Let x be even if and only if x^2 is even for all x in Z. Let there be a y in Z such that there is a z in Z such that x*y = z and there doesn't exist a d in Z such that d is a common divisor of y and z for all x in Q. Thus sqrt(2) isn't rational.*

Proof:

Let 2 be a common divisor of x and y if x is even and y is even for all y in Z for all x in Z. Let x be even if and only if x^2 is even for all x in Z. Let there be a y in Z such that there exists a z in Z such that x*y = z and there is no d in Z such that d is a common divisor of y and z for all x in Q.
We prove that sqrt(2) isn't rational by a contradiction. Let sqrt(2) be rational.

Let n in Z and let there be a dc_269 in Z such that sqrt(2)*n = dc_269 and there is no dc_273 in Z such that dc_273 is a common divisor of n and dc_269.
Let m in Z, let sqrt(2)*n = m and let there be no dc_279 in Z such that dc_279 is a common divisor of n and m.
We prove that m^2 = 2*n^2 in order to prove that there exists a dc_287 in Z such that m^2 = 2*dc_287. m^2 = 2*n^2 because sqrt(2)*n = m.
Hence m^2 is even. Hence m is even since m in Z. Then there exists a dc_343 in Z such that m = 2*dc_343.
Let k in Z and let m = 2*k. 2 in Z.
We prove that n^2 = 2*k^2 in order to prove that there is a dc_353 in Z such that n^2 = 2*dc_353. n^2 = 2*k^2 since m^2 = 2*n^2 and m = 2*k.
That implies that n^2 is even. That leads to even n because n in Z.
That leads to a contradiction because m in Z, n in Z, there is no dc_279 in Z such that dc_279 is a common divisor of n and m, m is even and 2 in Z.

Tactical TP in ΩMEGA

Result:

- 33 interactive steps
- resulting proof consists of 33 nodes
- expanded proof consists of about 200 nodes (automatic expansion)

Problematic (this also applies to other systems):

- tactics are not fitted to the problem at hand
- proving is tedious and user has to adapt to the system

Interactive islands planning

Declarative approach versus procedural approach

Interactive islands planning

Declarative approach versus procedural approach

Network of proof ‘islands’

$$\dfrac{\dfrac{2 * n^2 = m^2}{Even(m^2)}\,Island}{Even(m)}\,Island$$

$$\vdots$$

Interactive islands planning

Declarative approach versus procedural approach

Network of proof 'islands'

$$\cfrac{\cfrac{2*n^2=m^2}{Even(m^2)}\ Island}{Even(m)}\ Island$$

$$\vdots$$

- Islands structure the proof in natural form
- Islands provide no argument for soundness

$\Rightarrow$ Verification: expansion of island steps (automated, interactive, recursive island approach)

```
L6  (L4)       ! (INT N)                              ANDEL: (L4)
L8  (L5)       ! (INT M)                              ANDEL: (L5)
L11 (L5) ! (= (TIMES (SQRT 2) N) M)                   ANDE: (L9)
```

```
OMEGA: island-tactic
CONC (NDLINE) Conclusion of step: nil
PREMS (NDLINE-LIST) Premises of step: (L11 L6 L8)
PARAM (TERM) Formula of Conclusion: (= (times 2 (power n 2)) (power m 2))
```

```
L13 (L4 L5) |- (= (TIMES 2 (POWER N 2)) (POWER M 2)) ISLAND-TACTIC: (L11 L6 L8)
```

```
OMEGA: island-tactic nil (L13 L6 L8) (evenp (power m 2))
```

```
L14 (L4 L5) |- (EVENP (POWER M 2))                 ISLAND-TACTIC: (L13 L6 L8)
```

```
OMEGA: island-tactic nil (L14 L8) (evenp m)
```

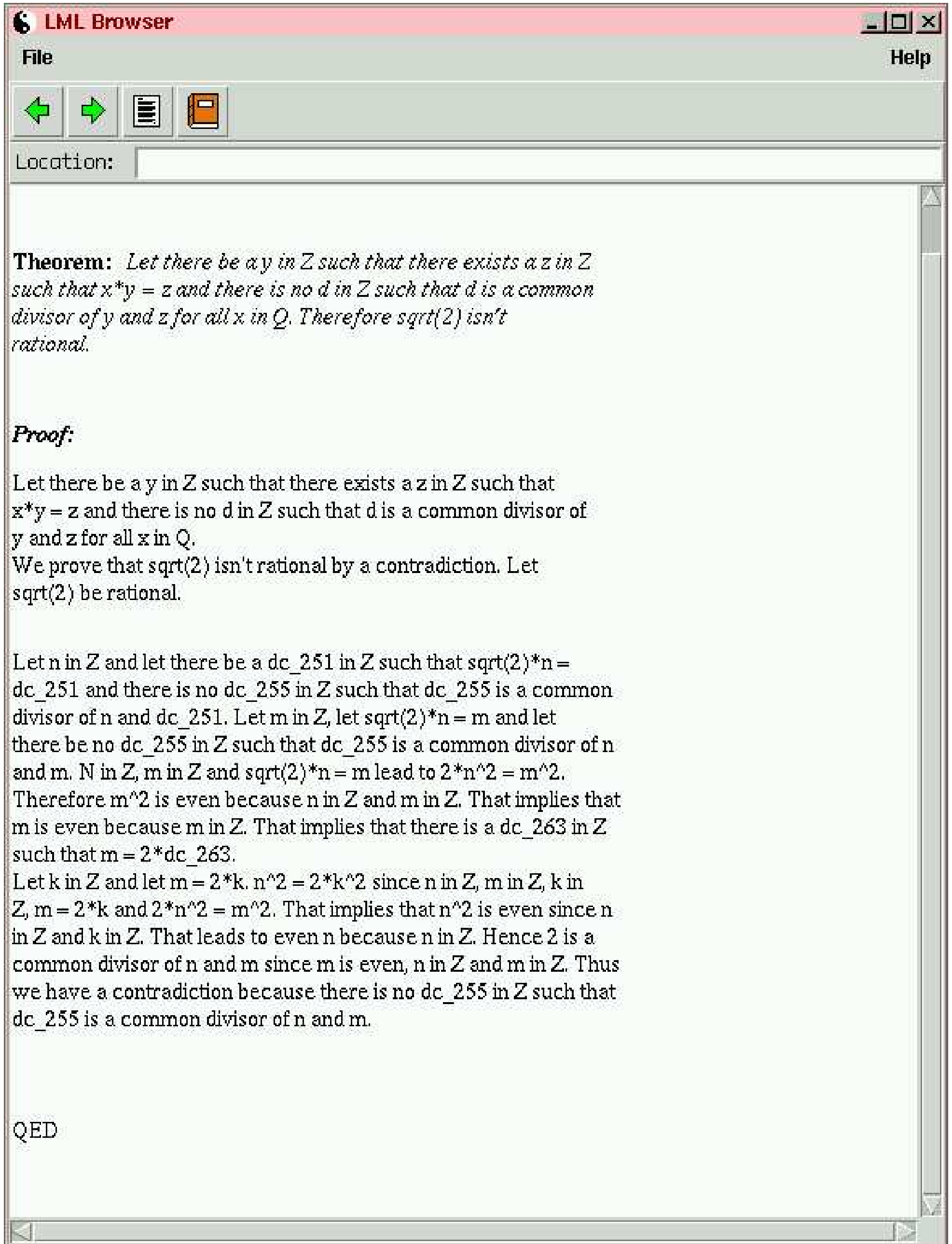
```
L15 (L4 L5) |- (EVENP M)                              ISLAND-TACTIC: (L14 L8)
```

Proof with Islands

Results:

- 15 interactive steps (8 island steps, 7 tactic steps)
- resulting proof consists of 25 nodes
- expanded proof consists of about 280 nodes (interactive expansion)

Proof with Islands



Theorem: *Let there be a y in Z such that there exists a z in Z such that x*y = z and there is no d in Z such that d is a common divisor of y and z for all x in Q. Therefore sqrt(2) isn't rational.*

Proof:

Let there be a y in Z such that there exists a z in Z such that x*y = z and there is no d in Z such that d is a common divisor of y and z for all x in Q.
We prove that sqrt(2) isn't rational by a contradiction. Let sqrt(2) be rational.

Let n in Z and let there be a dc_251 in Z such that sqrt(2)*n = dc_251 and there is no dc_255 in Z such that dc_255 is a common divisor of n and dc_251. Let m in Z, let sqrt(2)*n = m and let there be no dc_255 in Z such that dc_255 is a common divisor of n and m. N in Z, m in Z and sqrt(2)*n = m lead to 2*n^2 = m^2.
Therefore m^2 is even because n in Z and m in Z. That implies that m is even because m in Z. That implies that there is a dc_263 in Z such that m = 2*dc_263.
Let k in Z and let m = 2*k. n^2 = 2*k^2 since n in Z, m in Z, k in Z, m = 2*k and 2*n^2 = m^2. That implies that n^2 is even since n in Z and k in Z. That leads to even n because n in Z. Hence 2 is a common divisor of n and m since m is even, n in Z and m in Z. Thus we have a contradiction because there is no dc_255 in Z such that dc_255 is a common divisor of n and m.

QED

Summmary of Island Approach

- Sketch top-level proof in a declarative way
- In general: expansion of island steps generates proof object in its own right
- In our case study: Expansion of island steps with external systems almost completely automatic.
- Problems in the automatization:
  - Which defi nitions need to be folded or unfolded?
  - Which assumptions are missing?
  - Which facts need to be included from the database?

Proof Planning the Problem

Since ΩMEGA is also a proof planner: Can the proof can be automatically proof planned?

- First needed: Acquisition of methods by generalization of steps in island proof
- Possible then: automatically prove plan arbitrary problems of the $\sqrt[k]{l}$ is irrational domain
- But note: knowledge acquisition process is crucial for the success
- Methods still to problem specifi c: e.g. with respect to lemma retrieval or folding/unfolding of defi nitions
- See article submitted to Journal of Automated Reasoning

Proof Planning the Problem

(1) Use RAT-CRITERION and construct an indirect proof.

(2) To get a contradiction show that the two constants (existential variables) in RAT-CRITERION, which are supposed to have no common divisor, actually do have a common divisor d.

(3) To fi nd a common divisor transform equations (for example, $\sqrt{2} \cdot n = m \longrightarrow 2 \cdot n^2 = m^2$), derive new divisor statements (for example, from $2 \cdot n^2 = m^2$ derive that m^2 has divisor 2, or from m^2 has divisor 2 derive that m has divisor 2); derive from given divisor statements new representations of terms, and use them for further transformations.

Conclusion

Although the $\sqrt{2}$-example is mathematically trivial, it nevertheless provides a challenge for mathematical assistant systems: not about automation, but about "natural" interaction and proof construction.

The example particularly requires the combination of

- deduction
- computation
- lemma retrieval
- folding or unfolding definitions

There should be more such case studies!

Criteria for the comparison of systems?

Conclusion cont'd

ΩMEGA is much more than just a proof planner:

- tactical theorem prover
- new: interactive island planning
- it provides various integrated support tools

Automated proof planning of $\sqrt[k]{l}$-examples is of course possible: by generalizing and programming reasoning patterns

But note the price to be paid: knowledge acquisition!

There is still much to do! And the main problem is not that we need stronger "general" proof tools!

Future of OMEGA

- . . . of course, many things to mention . . .
- Central issue at the moment:
  - New Logic Layer based on Core System instead if ND
- Benefi ts:
  - Core hides many aspects of the logic layer from the user.
  - . . . many further group internal benefi ts . . .

Future of OMEGA

A1 *symmetric*(A)

A2 *symmetric*(B)

$\ldots$

G *symmetric*$(A \cap B)$

System suggestion: Apply-Assertion(def-symmetric) then you would get the following proof state . . .

Future of OMEGA

A1 *symmetric*(A)

A2 *symmetric*(B)

$\ldots$

G $\forall_{x,y} \langle x, y \rangle \in A \cap B \Rightarrow \langle y, x \rangle \in A \cap B$

System suggestion: Still Apply-Assertion(def-symmetric) ...

Future of OMEGA

A1 $\textit{symmetric}(A)$

A2 $\textit{symmetric}(B)$

A3 $\forall_{x,y}\langle x, y\rangle \in A \Rightarrow \langle y, x\rangle \in A$

A4 $\forall_{x,y}\langle x, y\rangle \in B \Rightarrow \langle y, x\rangle \in B$

$\ldots$

G $\forall_{x,y}\langle x, y\rangle \in A \cap B \Rightarrow \langle y, x\rangle \in A \cap B$

System suggestion: Focus on right-hand-side of implication . . .

Future of OMEGA

A1 *symmetric*(A)

A2 *symmetric*(B)

A3 $\forall_{x,y} \langle x, y \rangle \in A \Rightarrow \langle y, x \rangle \in A$

A4 $\forall_{x,y} \langle x, y \rangle \in B \Rightarrow \langle y, x \rangle \in B$

A4 $\langle c_1, c_2 \rangle \in A \cap B$

$\ldots$

G $\langle c_1, c_2 \rangle \in A \cap B$

System suggestion: Apply-Assertion(def-$\cap$) $\ldots$

Future of OMEGA

A1 $\textit{symmetric}(A)$

A2 $\textit{symmetric}(B)$

A3 $\forall_{x,y}\langle x,y\rangle \in A \Rightarrow \langle y,x\rangle \in A$

A4 $\forall_{x,y}\langle x,y\rangle \in B \Rightarrow \langle y,x\rangle \in B$

A4 $\langle c_1,c_2\rangle \in A \cap B$

A5 $\langle c_1,c_2\rangle \in A \wedge \langle c_1,c_2\rangle \in B$

$\ldots$

G $\langle c_2,c_1\rangle \in A \wedge \langle c_2,c_1\rangle \in B$

System suggestion: Apply-Assertions(A3,A4) . . .

Future of OMEGA

A1 *symmetric*(A)

A2 *symmetric*(B)

A3 $\forall_{x,y}\langle x,y\rangle \in A \Rightarrow \langle y,x\rangle \in A$

A4 $\forall_{x,y}\langle x,y\rangle \in B \Rightarrow \langle y,x\rangle \in B$

A4 $\langle c_1,c_2\rangle \in A \cap B$

A5 $\langle c_1,c_2\rangle \in A \wedge \langle c_1,c_2\rangle \in B$

A6 $\langle c_2,c_1\rangle \in A$

A7 $\langle c_2,c_1\rangle \in B$

G $\langle c_2,c_1\rangle \in A \wedge \langle c_2,c_1\rangle \in B$

System suggestion: Goal proved by Apply-Assertions(A6,A7) . . .

Future of OMEGA

1.	1;	$\vdash$	*symmetric*(A)	Hyp	
2.	2;	$\vdash$	*symmetric*(B)	Hyp	
3.	1,2;	$\vdash$	$\forall_{x,y} \langle x, y\rangle \in A \Rightarrow \langle y, x\rangle \in A$	AA[Sym-Def]	1
4.	1,2;	$\vdash$	$\forall_{x,y} \langle x, y\rangle \in B \Rightarrow \langle y, x\rangle \in B$	AA[Sym-Def]	2
5.	5;	$\vdash$	$\langle c_1, c_2\rangle \in A \wedge \langle c_1, c_2\rangle \in B$	Hyp	
6.	1,2,5;	$\vdash$	$\langle c_2, c_1\rangle \in A$	AA[3]	5
7.	1,2,5;	$\vdash$	$\langle c_2, c_1\rangle \in B$	AA[4]	5
8.	1,2,5;	$\vdash$	$\langle c_2, c_1\rangle \in A \wedge \langle c_2, c_1\rangle \in B$	AA[6,7]	6 7
9.	1,2;	$\vdash$	$\forall_{x,y} \langle x, y\rangle \in A \cap B \Rightarrow \langle y, x\rangle \in A \cap B$	AA[$\cap$-Def]	5 8
10.	1,2;	$\vdash$	*symmetric*$(A \cap B)$	AA[Sym-Def]	9