MathWiki
a Web-based Collaborative Authoring Environment for Formal Proofs

ICIS colloquium Radboud Universiteit Nijmegen
April 21 2008

Herman Geuvers

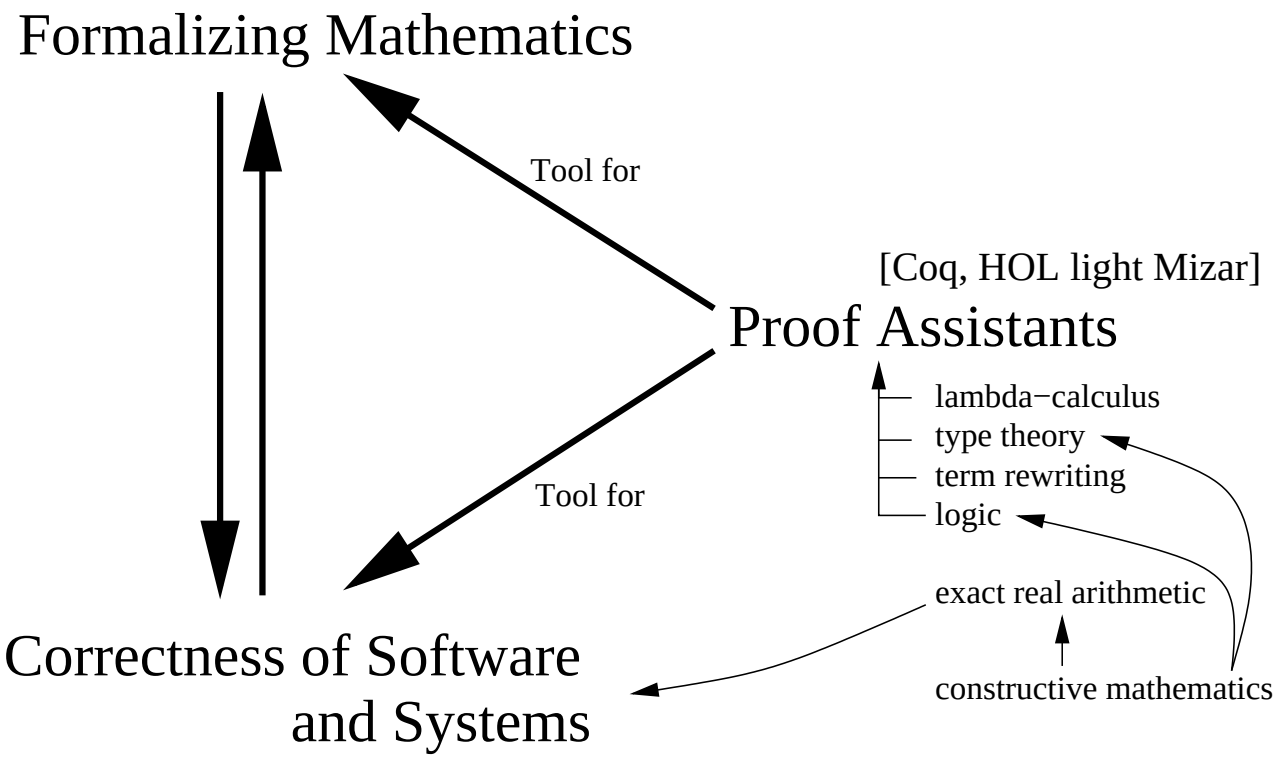
Joint work with Pierre Corbineau, Cezary Kaliszyk, James McKinna, Freek Wiedijk

MathWiki: EU project

MathWiki:
a Web-based Collaborative Authoring Environment for Formal Proofs
Application for a STREP project in EU FP7
Challenge 4: Digital Libraries and Content

- Radboud Universiteit Nijmegen
- Università di Bologna
- University of Edinburgh
- Technische Universität München
- INRIA Paris
- Uniwersytet w Białymstoku
- Jacobs University Bremen

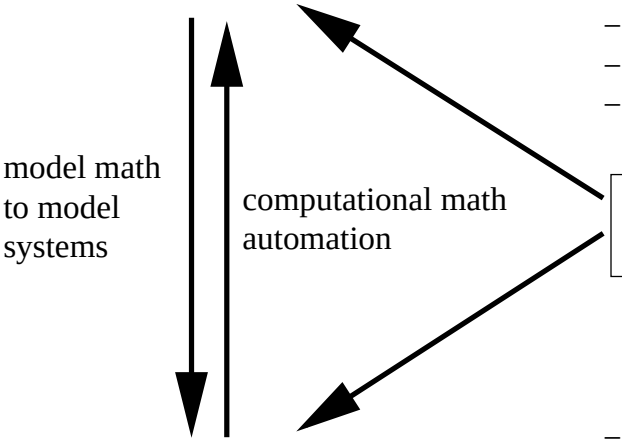
How this fits into the research of the PI group Foundations



Three Research themes and their interaction

Six Academic themes

Research Projects - PI group Foundations ICIS



- Web deduction
- Interactive Math. Docs
- MathWiki
- C−CoRN repository
- MathMode (declarative mode)
- Fear (complex analysis)

- Verification of Hybrid Systems
- Dependently typed programming
- Type systems & Static Analysis

Brain – Models of consciousness

Central topic: Systems for integrating proving and programming

One special research topic: Studying models of the brain.

Research Projects - PI group Foundations ICIS

- Web deduction: a web-based system for students to learn logic.
- Interactive Mathematical Documents: Integration of document editing and formalization of mathematics
- MathWiki: Wikipedia for formalized mathematics
- C-CoRN: Our library of constructive maths. formalized in Coq.
- MathMode: Declarative proof mode for Coq
- Fear: Formalizing equations in complex analysis.
- Verifying Hybrid Systems: model and verify hybrid systems in Coq
- Dependent types: programming in a richly typed language.

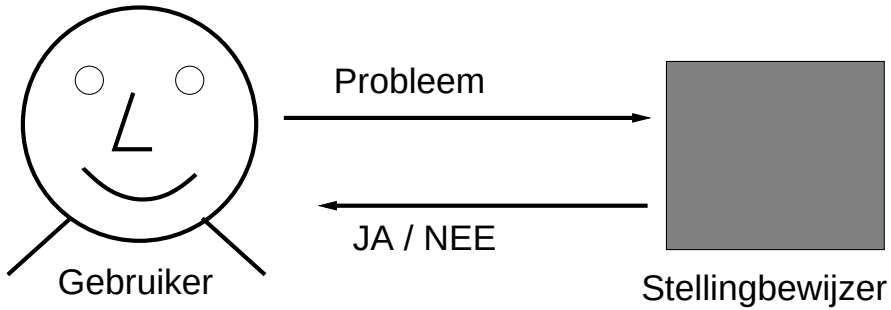
MathWiki

- Background and motivations
- Vision
- (Technical) Issues

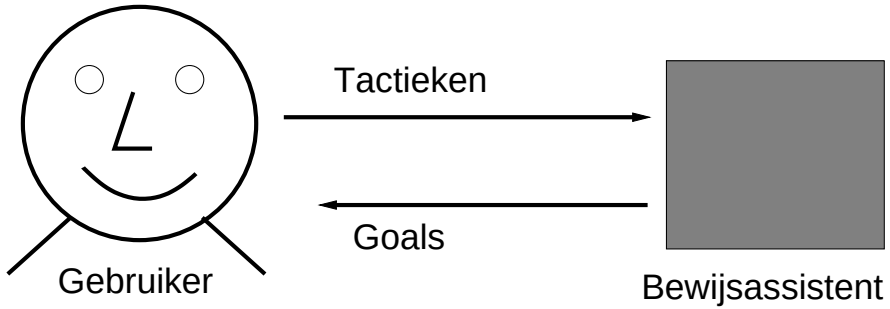
Background and motivations

Proof Assistants

- Theorem Prover? Automatic?



- Proof Assistant: Interactive!



Proof Assistants. Some Claims

- Claim 1 PAs are useful for modelling and verification of systems
- Claim 2 A formal representation is useful to communicate and really understand all the details of the mathematics
- Claim 3 We can extract/generate readable mathematical documents from a formalisation.
- Claim 4 PAs are useful for teaching logic and mathematics.

Proof Assistants. Some Claims

- Claim 1 + PAs are useful for modelling and verification of systems
- Claim 2 + A formal representation is useful to communicate and really understand all the details of the mathematics
- Claim 3 +/- We can extract/generate readable mathematical documents from a formalisation.
- Claim 4 +/- PAs are useful for teaching logic and mathematics.

Using Proof Assistants

For PA systems to be really useful for mathematical users we need

1. More automation. Things that are mathematically easy [according to a user] should be easy for the PA.
2. A less system dependent notation and way of interaction. Less verbose, less idiosyncratic.
3. Large, useable library of known results. Things that a mathematical user expects to be available should be available and possible to find.

We want to focus on 3 (and 2).

Message: There is not enough formalised mathematics

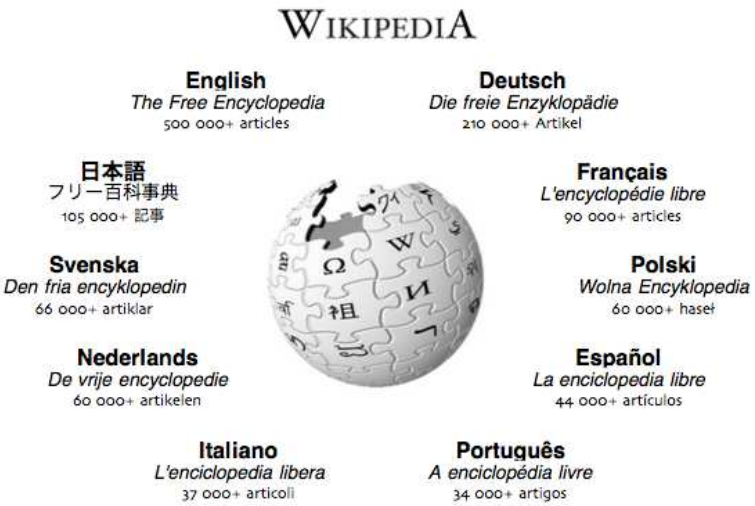
More mathematics needs to be formalised!

To formalise the undergraduate program of mathematics requires 140 man year. [Freek Wiedijk]

One research group will not be able to do this.

Solution: let the whole world participate to create a shared repository of formalised mathematics.

Wikipedia



A joint distributed development of a coherent on-line encyclopedia.

"Doesn't work in theory . . . but works in practice"

Vision

Aim Our aim is to open up to a wider community the rich collections of knowledge stored in the repositories of proof assistants and to facilitate the extension and editing of these repositories by outside users.

The further reaching aim is to forward the use of computer formalized mathematics and to establish the medium of computer checkable formal proofs as a valuable asset in ICT, notably in verification and correctness of software and systems.

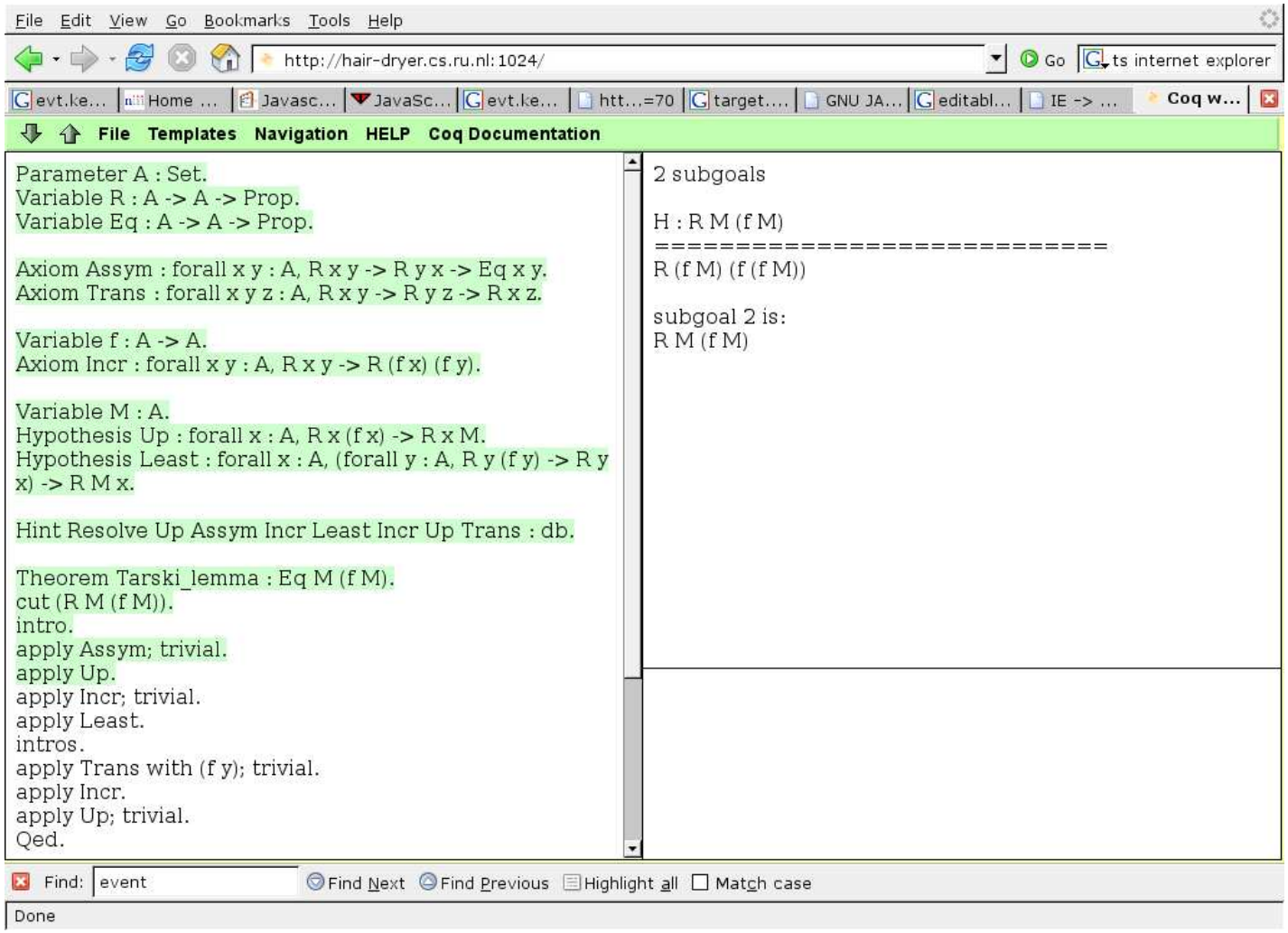
Wikipedia for PA repositories

- Claim 1 The Wikipedia approach also works for semantically rich (very structured) data.
  Consistency Issues!
- Claim 2 We can create attractive, useful web-pages for mathematical notions with formal content.

MathWiki: Two views

- "Bottom up" (PA technology push)
  - Support for joint distributed formalization (through a web interface)
  - Support for creating cross links (between
  - Search and High level presentation of content

Formalization through a web interface

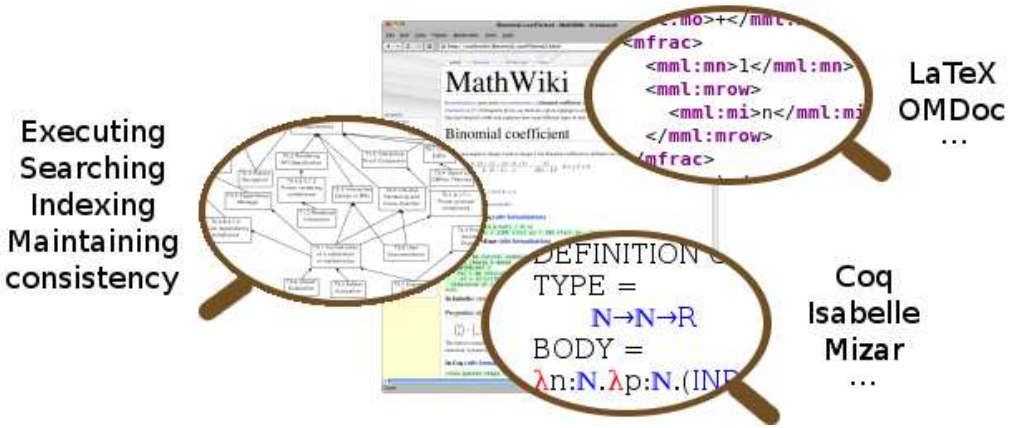


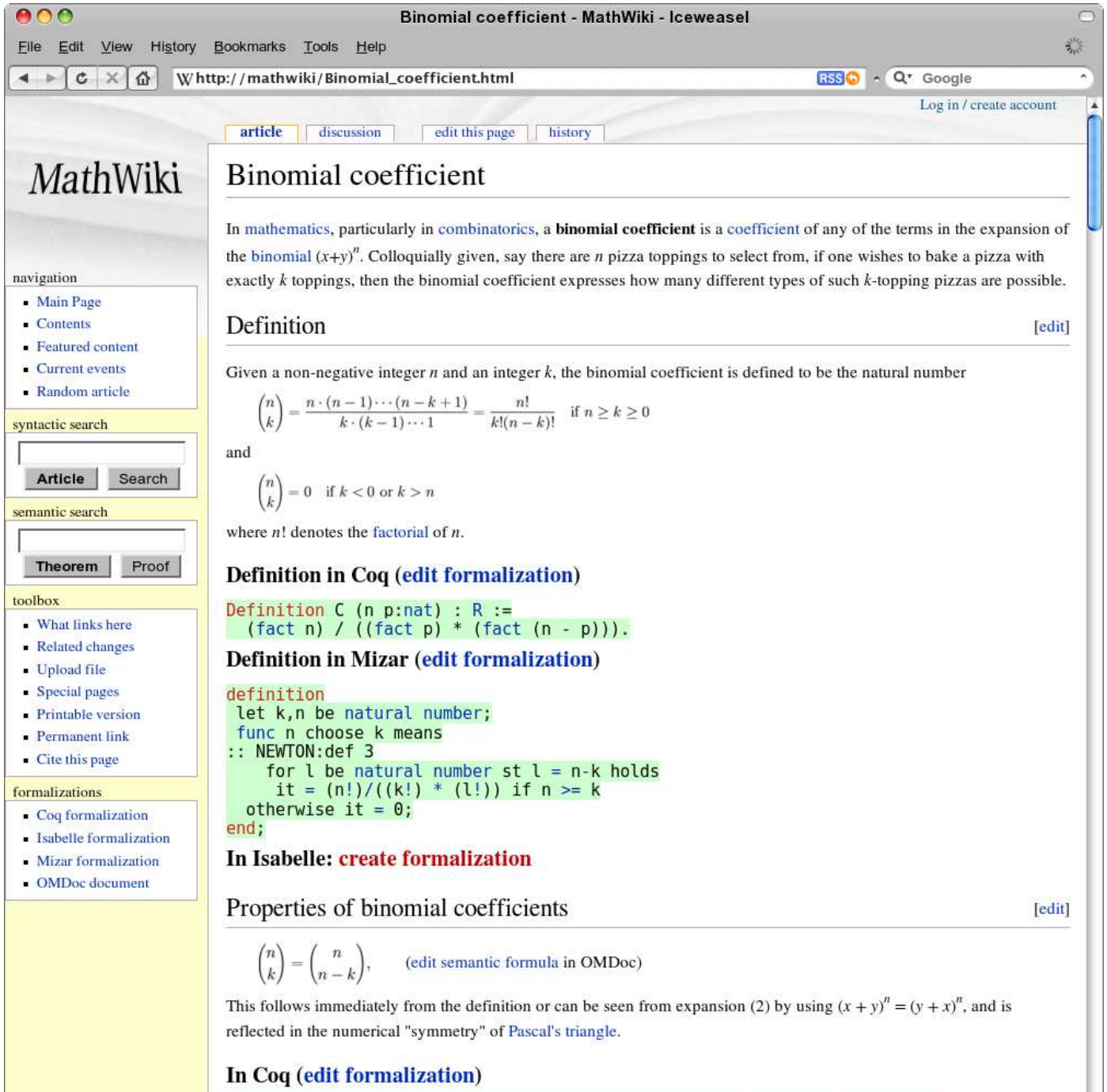
MathWiki: Two views

- "Top down" (Math communication pull)
  - Present one page for a mathematical notion, with (some) formal content and links.
  - Support for creating high level pages plus links to formal content
  - Compete with Wikipedia, MathWorld, ...

An example MathWiki page: binomial coefficient

Logo:




Binomial coefficient - MathWiki - Iceweasel
File Edit View History Bookmarks Tools Help
http://mathwiki/Binomial_coefficient.html
Google
Log in / create account
article discussion edit this page history
MathWiki
navigation
Main Page
Contents
Featured content
Current events
Random article
syntactic search
Article Search
semantic search
Theorem Proof
toolbox
What links here
Related changes
Upload file
Special pages
Printable version
Permanent link
Cite this page
formalizations
Coq formalization
Isabelle formalization
Mizar formalization
OMDoc document
Binomial coefficient
In mathematics, particularly in combinatorics, a binomial coefficient is a coefficient of any of the terms in the expansion of the binomial $(x+y)^n$. Colloquially given, say there are n pizza toppings to select from, if one wishes to bake a pizza with exactly k toppings, then the binomial coefficient expresses how many different types of such k-topping pizzas are possible.
Definition [edit]
Given a non-negative integer n and an integer k, the binomial coefficient is defined to be the natural number
$\binom{n}{k} = \frac{n\cdot(n-1)\cdots(n-k+1)}{k\cdot(k-1)\cdots 1} = \frac{n!}{k!(n-k)!}$ if $n \geq k \geq 0$
and
$\binom{n}{k} = 0$ if $k < 0$ or $k > n$
where $n!$ denotes the factorial of n.
Definition in Coq (edit formalization)
Definition C (n p:nat) : R :=
(fact n) / ((fact p) * (fact (n - p))).
Definition in Mizar (edit formalization)
definition
let k,n be natural number;
func n choose k means
:: NEWTON:def 3
for l be natural number st l = n-k holds
it = (n!)/((k!) * (l!)) if n >= k
otherwise it = 0;
end;
In Isabelle: create formalization
Properties of binomial coefficients [edit]
$\binom{n}{k} = \binom{n}{n-k},$ (edit semantic formula in OMDoc)
This follows immediately from the definition or can be seen from expansion (2) by using $(x+y)^n = (y+x)^n$, and is reflected in the numerical "symmetry" of Pascal's triangle.
In Coq (edit formalization)

What are the selling points?

- The potential users are:
  - Expert users of PAs (computer scientists, engineers, verification, modelling, ...)
  - People interested in a precise mathematical description / proof. On the top level, it should be readable for undergraduates, without any knowledge of the PA

What are the selling points?

- Emphasis is on libraries; the repository is not a loose collection of individual contributions, but a documented coherent library of formalized mathematics.
Also: documentation of the prover itself, reference manual, tutorial

- High level access to precise formal mathematics.
A search for a mathematical concept should produce one page, with some formal content and links to the formal details.

What are the selling points?

- No local installation of a PA, always the latest version, no version management

Technical and other issues

A Web-based Collaborative Authoring Environment for Formal Proofs: What to do /develop?

- Collaborative environment for repositories (a semantic wiki)
- Web-based interface for various PAs
- Consistency management for the repositories (version management "plus")
- Search accross repositories and high level pages.
- Generic cross-system ontologies and metadata

Remarks on database (repository) management

- Concurrent access / Dependency analysis
- Consistency check (update crawler): saving not allowed if not consistent.
- History navigation (Older states of the repository)
- Support for import of large data sets (existing repostories), and export.

Remarks on semantic aspects

- Metadata /ontology (version number, author, dependencies, cross links between repositories, outside links, notation, . . .)
- Search for "similar" concepts?
- Proof development by "stepwise refinement"
- Generic high level proof language ??
- Formal translations between systems (in stead of cross links)??

Validation

- Import of existing developments
- Doing a new (large, joint) formal development
- Creation of content MathWiki pages for a specific mathematical theme. (E.g. real analysis.)
- End User Panel
- Challenge problems / Proof ideas / Proof Sketches / Prize puzzles / links to JFR
- Impact evaluation (compare with MathWorld, Wikipedia, . . .)

Some content issues w.r.t. MathWiki as a EU project

- Which PAs to include?
  Start with Coq, Isabelle, Mizar. Open to other systems.
- Open standards, open source.
- Which functionality is joint for more PAs? (Or can be made to be joint?)
- File ownership? No
- Library committee? Yes

Expertise of the partners

- Radboud Universiteit Nijmegen (Web interfaces, PA repositories, Coq, Mizar)
- Università di Bologna (PAs, Search, Metadata)
- University of Edinburgh (Prover Interfaces)
- Technische Universität München (Isabelle)
- INRIA Paris (Coq)
- Uniwersytet w Białymstoku (Mizar)
- Jacobs University Bremen (OMDoc, ontologies, semantic web)