**Task 1. Answer Retrieval**
Given a Math Question, Find Answers

**Answer Retrieval (Task 1)**

**Question (Topic A.2.39)** Let $R$ be a symmetric, transitive relation. If $(x, y)$ then the symmetric property implies that $(y, z)$. Using the transitive property upon $(x, y)$ and $(y, z)$ we can conclude $(x, z)$. Is this fair logic or is it flawed?

**RELEVANT**
Symmetric means we can represent the relation as an undirected graph. Transitive means this graph is composed of connected components which either look like a point or $K_n$ with each point connected to itself. Reflexivity means each point is connected to itself. Thus, a necessary and sufficient condition is that one component is not a point (i.e., one element is not related to any other). In exactly these cases your proof fails because it requires each $x$ to be related to another $y$.

**NON-RELEVANT**
You should start with $(x, y) \in R$ and $(y, z) \in R$ and from these two statements show that $(x, z) \in R$.
This shows transitivity. Now, the two statements show that indeed $(x, z) \in R$ by definition (using $y$ as intermediate).

**Task 2. Formula Retrieval**
Given a Math Formula, Find Formulas

**Formula Retrieval (Task 2)**

**FORMULA QUERY (Topic B.2.20)** I’m having a difficult time understanding how to give a combinatorics proof of the identity

$$\sum_{k=0}^{n} \binom{x + k}{k} = \binom{x + n + 1}{n}$$

**RELEVANT**
Question: prove by induction on $n+m$ the combinatoric identity:

$$\sum_{k=0}^{n} \binom{m + k}{k} = \binom{m + n + 1}{n}$$

I’ve tried to do on both $n$ and $m$.

**Hint**

$$\sum_{k=0}^{n} \binom{n}{k} x^k = (1 + x)^n$$

Integrate twice both rhs and lhs with respect to $x$ and when finished, plug $x = 1$ in your result.

**Task 3. Question Answering (NEW!)**
Given a Math Question, Generate Answers

**Question Answering (Task 3)**

**QUESTION (Topic A.1.99)** Is it correct to say that this set $E = (0, 1]$ where $E \subseteq R$ (Where $R$ is the set of real numbers) is not closed?

**RELEVANT**
Yes. Closed sets are the complements of open sets. An open set is a set for all points there exists a neighborhood contained in the set with positive radius.

**NON-RELEVANT**
think this method can be reduced using three coefficients (1, 3, 2) to know if a number is a multiple of seven or not. We multiply the last number by 1, the second from the right by 3, and finally by 2. Then next three digits by (1, 3, 2) and again by positive.

**Test Collection**
Posts from Math Stack Exchange (2010-2018)
Formulas Index with LaTeX and MathML Representations

**Task 1/3 Topics:** 100 Questions from Posts in 2021
**Task 2 Topics:** 100 Formulas from Questions in Task 1

**Evaluation**
**Task 1:** Top-1000 answers for each question
**Task 2:** Top-1000 formulas for each query
**Task 3:** 5 answers for each question

NDCG’ as primary measure, with $P’@10$ and MAP’
De-duplication for Task 2 with Visually distinct formulas

**Available Tools**
Baseline: **Terrier** (text search) and **Tangent-S** (formula search)
Training data for Tasks 1 and 2 from ARQMath-1 and -2
Runs and Systems from ARQMath-1 and -2

---

#ARQMath
Behroz Mansouri, Vítek Novotný, Anurag Agarwal, Douglas W. Oard, and Richard Zanibbi
{bm3302, axasma, rxzvcs}@rit.edu, witiko@mail.muni.cz, oard@umd.edu