University of Illinois at Urbana-Champaign
Department of Computer Science

Midterm Solutions

CS 273 Introduction to Theoretical Computer Science
Spring, 2002

Multiple Choice Problems, 2 points each

1. On what variable would you induce to prove the following identity?

$$\sum_{k=0}^{m}\left(\binom{m-2}{k-1}+\sum_{l=k-2}^{k}\binom{m-2}{l}\right)=2^m$$

   Answer: m

2. How many ways are there to arrange the letters of the word NONSENSE?
   Answer: $\frac{8!}{3!\cdot 2!\cdot 2!\cdot 1!}$

3. How many integers are there from 1 to 1000 *not* divisible by 2,3, or 10?
   Answer: $333 = 1000-(500+333+100-166-33-50+16)$ by Inclusion-Exclusion Formula

4. How many bit strings of length 8 are there that contain a pair of consecutive 0s?
   Answer: $201 = 2^8 - F_{8+2} = 256 - 55$.

5. How many n variable functions are there where each of the variables has domain of size m and the function range is of size r?
   Answer: r^{m^n}

6. What is $\sum_{k=0}^{n} 2^{-k}\binom{n}{k}$?
   Answer: $3^n/2^n = (1+1/2)^n = (1+2^{-1})^n$

7. There are 10 lockers in the athletic center and 76 students who need lockers. Therefore, some students must share lockers. What is the largest number of students who must necessarily share a locker?
   Answer: $8 = \lceil 76/10 \rceil$ by pigeonhole principle

8. What is the number of integral solutions to $x+y+z=32$, where $x,y,z\geq 0$?
   Answer: $\binom{34}{2}$

9. What is the recurrence for the number of ways to climb a staircase, if you go up either 1 or 3 steps at a time?
   Answer: $a_n = a_{n-1} + a_{n-3}$

10. How many initial values will you need to calculate in order to solve the following recurrence exactly using annihilators?

$$a_n = 4a_{n-1} - 6a_{n-2} + 4a_{n-3} - a_{n-4} + 2^n + \phi^n + (2\cdot\phi)^n + (2+\phi)^n?$$

   Answer: 8 since the annihilator is $(E^4+4E^3+6E^2-4E+1)(E-2)(E-\Phi)(E-2\Phi)(E-(2+\Phi))$

Short Problems, 5 points each, all or nothing (no justification necessary)

1. What is $\sum_{k=0}^{1{,}000{,}000} k$?
   Answer:
   $$\sum_{k=0}^{1{,}000{,}000} k = \frac{1,000,000 \cdot 1,000,001}{2} = 500,000,500,000$$

2. What is $\binom{6,000,000,001}{3}$?
   Answer:
   $$\begin{aligned}\binom{6,000,000,001}{3} &= \frac{6,000,000,001 \cdot 6,000,000,000 \cdot 5,999,999,999}{6} \\ &= 10^9(6 \cdot 10^9 + 1)(6 \cdot 10^9 - 1) \\ &= 10^9(36 \cdot 10^18 - 1) \\ & 35,999,999,999,999,999,999,000,000,000\end{aligned}$$

3. What is the coefficient of x^{10} in the expansion of $\left(x + \frac{1}{x}\right)^{100}$?
   Answer: This is the term $\binom{100}{k} x^k (1/x)^{100-k}$ where $k-(100-k) = 10$. Therefore $k = 110/2 = 55$. So the coefficient of x^10 is $\binom{100}{55} = \binom{100}{45}$.

4. Three married couples have purchased 6 seats in a row for a concert. In how many different ways can they be seated left to right if:
   (a) There are no restrictions on seating?
   Answer: $6! = 720$
   (b) If each couple must sit together?
   Answer: $3! \cdot 2^3 = 48$
   (c) If all men sit together to the right of all women?
   Answer: $3!3! = 36$
   (d) If the seating arrangement must be "boy-girl"?
   Answer: $3!3! = 36$
   (e) If the seating arrangement must be "boy-girl", and each couple must sit together?
   Answer: $3! = 6$

Long Problems

1. **Combinatorics**

   Prove the following in two ways
   $$\binom{2n}{2} = 2\binom{n}{2} + n^2$$

(a) algebraically (but without induction)

Solution: There are two general ways to do this, one by translating directly to factorials, the other using binomial identities.

The first way just simplifies both sides using factorials and elementary algebra. First the left hand side (LHS)

$$\begin{aligned}
\binom{2n}{2} &= \frac{(2n)!}{(2n-2)!2!} & \binom{n}{m} = \frac{n!}{(n-m)!m!} \\
&= \frac{2n(2n-1)}{2} & \text{divide out like factors} \\
&= n(2n-1) \\
&= 2n^2 - n
\end{aligned}$$

Now for the right hand side (RHS):

$$\begin{aligned}
2\binom{n}{2} + n^2 &= 2\frac{n!}{(n-2)!2!} + n^2 & \binom{n}{m} = \frac{n!}{(n-m)!m!} \\
&= n(n-1) + n^2 & \text{divide out like factors} \\
&= n^2 - n + n^2 \\
&= 2n^2 - n
\end{aligned}$$

Since both the LHS and the RHS simplify to identical formulas (through equality preserving transformations), they are equal. This is probably the best way of finding the proof. One can also rewrite it one linear derivation by reversing the second simplification and appending it to the first:

$$\begin{aligned}
\binom{2n}{2} &= \frac{(2n)!}{(2n-2)!2!} \\
&= \frac{2n(2n-1)}{2} \\
&= n(2n-1) \\
&= 2n^2 - n \\
&= n^2 - n + n^2 \\
&= n(n-1) + n^2 \\
&= 2n(n-1)/2 + n^2 \\
&= 2\frac{n!}{(n-2)!2!} + n^2 \\
&= 2\binom{n}{2} + n^2
\end{aligned}$$

This is a perfectly acceptable proof because all the manipulations are legal. And you don't have to remember a previous result (well, you hope you write everything down anyway), because you went in one direct line from LHS to RHS. This sort of proof has the tendency to baffle the reader because they can so easily verify it, but they have absolutely no idea where the magic came from for lines 5 and 7. Why split the n^2? Why multiply by 2 just to divide it out? It looks like you're going backwards! I'm not saying this is an inappropriate way of expressing a proof, you should just be aware of your proof method and the expression of your proof and its affects on the reader.

Common mistakes for the above method are:

- mixing up sides when doing the derivation side by side. Doing a side by side derivation (that is something like $\binom{2n}{2} = 2\binom{n}{2} + n^2 \Rightarrow \ldots \Rightarrow \ldots \Rightarrow n(2n-1) = n^2 - n + n^2 \Rightarrow 2n^2 - n = 2n^2 - n$) is readable as a proof (and if all the operations are performed correctly) but is technically not really correct, because you don't really yet know if one side actually does equal the other. Also it can lead to egregious errors like assuming the thing you are trying to prove if you accidentally look at a previous line and use it as an equality, when it is only a thing to be proved.

- minor transcription errors, e.g. accidentally transferring $(2n)!$ on one line to $2n!$ on another (treating it as $2(n!)$) (then you probably made a second mistake later to make everything turn out ok in the end.

The second way of proving the identity is by using Vandermonde's identity with $m = n$ and $r = 2$. This gives:

$$\begin{array}{rcll} \binom{2n}{2} & = & \binom{n+n}{2} & \\ & = & \sum_{k=0}^{2} \binom{n}{k}\binom{n}{2-k} & \text{Vandermonde's identity} \\ & = & \binom{n}{0}\binom{n}{2} + \binom{n}{1}\binom{n}{1} + \binom{n}{2}\binom{n}{0} & \text{expanding the sum} \\ & = & \binom{n}{2} + n \cdot n + \binom{n}{2} & \text{small cases of binomials} \\ & = & 2\binom{n}{2} + n^2 & \text{algebra} \end{array}$$

This I would consider a clever method because with a little extra work (trying to find one of a vast sea of identities to one small problem) gives you a pretty short proof.

(b) combinatorially

Solution: First a few examples of what a combinatorial proof is *not*

- using words instead of symbols
- handwaving
- a picture proof

A combinatorial proof (of an identity) tries to give some interpretation of the two sides such that they count the same set of objects. You must come up with a counting situation S, interpret the LHS, showing that it counts A, then do the same for the RHS. Often one of these steps is trivial or obvious.

So here is a combinatorial proof:

Let S be the number of subsets of size 2 from the union of 2 (disjoint) sets of size n. The LHS counts this directly (this was the obvious interpretation given the description of A): the whole set is of size $2n$ and we want the number of subsets of size 2 (number of unordered pairs).

The RHS counts things by considering from which of the 2 sets (let's call the two sets A and B) does each item in the pair come. We'll look at the cases separately:

- Both items of the pair could come from A, there are $\binom{n}{2}$ ways to do that.
- Both items of the pair could come from B, there are $\binom{n}{2}$ ways to do that.
- One item could come from A and the other from B. There are n ways for the first and n ways for the second, for a total n^2

These three cases are mutually exclusive *and* they cover all the possibilities, so all we have to do is add them up to get $2\binom{n}{2} + n^2$.

Since we've counted the same situation in two different ways, the two different ways must be equal. QED.

This proof is not too long and tells you quite a bit more about the symbols than the algebraic proof. The difficulty is in finding such a proof. Often you can look at the symbols and translate to combinatorial situations you know already. All subsets can come from 2^n, subsets of a particular size come from $\binom{n}{k}$, disjoint sets from addition (the sum rule). A combinatorial proof asks a bit more creativity from you, but the insight you get from it is infinitely more than an algebraic proof (I claim that there is very little to no insight from an algebraic proof).

The common pitfalls in coming up with a combinatorial proof are

- using a binomial identity in an algebraic manner. A combinatorial identity is an algebraic formula (with equality) that describes a combinatorial situation. It may have been proved by combinatorics, algebra, calculus, whatever, and it may be interpreted combinatorially, algebraically, over the real or complex numbers, whatever. But if you use it in a proof in an algebraic manner, you are not using it combinatorially.
- just translating symbols into words. There is (usually) no proof in such a direct translation. The proof is in providing a correspondence between the situations that these words describe.
- picking the wrong interpretation. Sometimes the symbols can lead you astray. For example, the n^2, standing by itself, seems like an *ordered* pair *with* replacement, from *one* set. That is a reasonable starting point for n^2 but it might make the interpretation of the other symbols more complicated (how to reconcile 2 kinds of unordered pairs and 1 kind of ordered pair on a set of size n with an unordered pair of $2n$).

2. **Induction**

Prove the following by induction

$$\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$$

Solution: *Base case:* $n = 0$.

$$\begin{aligned} LHS &= \sum_{k=0}^{0} k\binom{n}{k} \\ &= 0. \end{aligned}$$

$$\begin{aligned} RHS &= 02^{0-1} \\ &= 0. \end{aligned}$$

Since LHS matches RHS, the statement holds for $n = 0$.

Induction Hypothesis: Let us assume that the statement holds for some $n \geq 0$ i.e. $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$ for some $n \geq 0$.

To Prove: $\sum_{k=0}^{n+1} k\binom{n+1}{k} = (n+1)2^n$.

$$\begin{aligned} LHS &= \sum_{k=0}^{n+1} k\binom{n+1}{k} \\ &= \sum_{k=0}^{n+1} k\left[\binom{n}{k} + \binom{n}{k-1}\right], \text{ using Pascal's identity} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n+1} k\binom{n}{k} + \sum_{k=0}^{n+1} k\binom{n}{k-1} \\
&= \sum_{k=0}^{n} k\binom{n}{k} + \sum_{k=0}^{n+1} k\binom{n}{k-1}, \text{since } \binom{n}{n+1} = 0 \\
&= n2^{n-1} + \sum_{k=0}^{n+1} k\binom{n}{k-1}, \text{using induction hypothesis} \\
&= n2^{n-1} + \sum_{j=-1}^{n} (j+1)\binom{n}{j}, \text{by substituting } j = k-1 \\
&= n2^{n-1} + \sum_{j=0}^{n} (j+1)\binom{n}{j}, \text{since } \binom{n}{-1} = 0 \\
&= n2^{n-1} + \sum_{j=0}^{n} j\binom{n}{j} + \sum_{j=0}^{n} \binom{n}{j} \\
&= n2^{n-1} + n2^{n-1} + \sum_{j=0}^{n} \binom{n}{j}, \text{by using induction hypothesis} \\
&= n2^{n} + \sum_{j=0}^{n} \binom{n}{j} \\
&= n2^{n} + 2^{n}, \text{by using the identity, } \sum_{j=0}^{n} \binom{n}{j} = 2^{n} \\
&= (n+1)2^{n} \\
&= RHS
\end{aligned}
$$

Hence the statement holds for $n+1$ also. Therefore by the principle of Mathematical Induction, the statement holds for all $n \geq 0$.

Some common mistakes were:

- Forgetting the base case
- Getting the summation indices wrong
- Proving the identity by algebraic methods without using induction

3. **Probability**

On the planet Bayleen, in the star system Tau Ceti, there is a species of life form that has three sexes: the ximander (with 20 eyes), the yeoman (with 30 eyes), and the zyzygy (with 500 eyes).

(a) How many children must be born before we can be sure that there are at least 17 of some sex? Why?

Solution: 5 points We want the number of children, say N, who must be born before we can be sure that there are at least 17 children of the same sex. Let us find out the maximum number of children that can be born without satisfying this condition. We can have 16 children of each of the sexes, and yet not have 17 of any one sex. This number is 48. Therefore, if we have one child more, it is bound to be from one of the three sexes, and hence, you will have at least 17 of any one sex. Therefore, N is $16 * 3 + 1 = 49$.

This also follows directly from the pigeonhole principle. We want to find the smallest integer N such that $\lceil N/3 \rceil = 17$. Since $(3 * 16)/3 = 16, \lceil (3 * 16 + 1)/3 \rceil = 17$, so $N = 3 * 16 + 1 = 49$.

Regardless of the wishes of the parents, there is always a 60% chance that a ximander will be born, a 30% chance for a yeoman, and a 10% chance for a zyzygy.

(b) For a group of three children, what is the probability that no two have the same sex?

Solution: 3 points

The probability that no two of them have the same sex would be the probability that we have one child from each of the sexes.

Let us consider an equivalent situation. Suppose, we have $10n$ elements total comprised of three different types of elements. The number of elements of the first type is $6n$, the number of elements of the second type is $3n$ and the number of elements of the third type is n. The problem is to pick out one element of each type. Now, the probability of choosing 3 elements each of a different type from $10n$ elements is $\frac{\binom{6n}{1}\binom{3n}{1}\binom{n}{1}}{\binom{10n}{3}}$. When n tends to infinity, this value becomes $6 \cdot (0.6) \cdot (0.3) \cdot (0.1)$. This is essentially the answer to the question of having each child of a different sex in a group of three children. Perhaps the more intuitive way to proceed would have been to consider the probability of not having each child in the group have a different sex, and then subtract that from one, which would work out to the same value. Hence, the probability of having each child from a different sex in a group of three would be $6 \cdot (0.6) \cdot (0.3) \cdot (0.1)$.

(c) What is the expected number of eyes a child will have?

3 points We know that expectation is given by $\sum_{i=0}^{n} x_i \cdot p(x_i)$. Here, this evaluates to $(0.6) \cdot (20) + (0.3) \cdot (30) + (0.1) \cdot (500) = 12 + 9 + 50 = 71$. Therefore, the expected number of eyes a child will have is 71. Of course, this does not mean that any single child will have 71 eyes. It only means that if we have n children, then the total number of eyes they would have would be $71n$ in the limit.

(d) How many children can parents expect to have before getting a yeoman?

4 points Suppose the parents have a yeoman as their first child, the probability of this happening is 0.3. Now, the number of children the parents had before the yeoman is 0. Now let us consider the case when the parents have a yeoman as their second child, the probability of this happening is (0.7)(0.3) and the number of children they had before having a yeoman is 1. Similarly, we can find out the probability of their having a yeoman as their third child and so on.

Therefore, expectation is $0 * (0.3) + 1(0.7)(0.3) + 2(0.7)^2(0.3) + 3(0.7)^2(0.3) + ...$

This is a geometric distribution of the form $p(1-p)^k$, the mean of which is $\frac{(1-p)}{p}$ or $\frac{1}{p} - 1$. Solving using the above, we see that the expected number of children before having a yeoman is 2.33.

Another way of looking at it would be if we consider the number of children to be inclusive of the yeoman. Thus, the probability of having a yeoman would be (0.3) and the number of children would be 1. If the parents had two children (second of which would be the yeoman) the probability of that would be $(0.7) * (0.3)$. We can calculate the probabilities of having three children in the same way and so on.

Therefore, expectation would be $1 * (0.3) + 2 * (0.7)(0.3) + 3(0.7)^2(0.3) + ...$

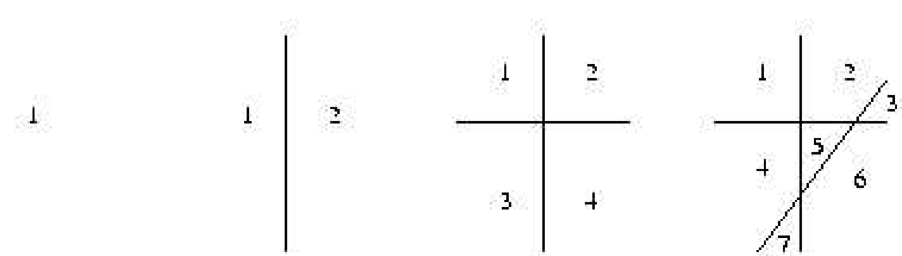
This series is the geometric distribution of the form $p(1-p)^{k-1}$ with mean $\frac{1}{p}$. On solving, we see that the expected number of children is 3.33. But this number includes the child which is a yeoman, so the number of children born before the yeoman would be $3.33 - 1 = 2.33$.

4. **Recurrence**

Consider a collection of n lines in the plane, such that each line intersects every other line and no three lines meet at a point.

(a) Let r_n be the number of regions into which n intersecting lines cut the plane. What are r_0, r_1, r_2, r_3?

Solution: 1 point $r_0 = 1$ No lines leave the whole plane as one region
$r_1 = 2$ One line cuts the plane into two regions
$r_2 = 4$ Two non-parallel lines cut the plane into four regions
$r_3 = 7$ The most three lines can cut the plane into is seven regions



Common mistakes:

- Not realizing the whole plane is one region
- Drawing parallel lines

(b) Obtain a recurrence for r_n. Justify.
Solution: 4 points Let there be $n-1$ lines on the plane cutting it into r_{n-1} regions. When drawing a new line, it has to intersect all of those $n-1$ old lines. The new line passes through a region, cutting it into two, before each intersection point. For $n-1$ lines there are $n-1$ intersection points and therefore $n-1$ new regions added before each intersection. In addition, there is one more region added after the last intersection point. Thus the total, there are n new regions in addition to the old r_{n-1}. So the recurrence

$$r_n = r_{n-1} + n$$

Another way to *write* this recurrence is

$$r_n = 2r_{n-1} - r_{n-2} + 1$$

But there is no combinatorial justification for this, it just comes out the same algebraically.
Common mistakes:

- Justifying the recurrence by "I see the pattern in the sequence..."
- Justifying the recurrence by "A new line adds n new regions" without saying why there are n new regions

(c) Solve the recurrence for r_n to obtain a formula for the number of regions.
Solution: 3 points

$$r_n = r_{n-1} + n, r_0 = 1, r_1 = 2, r_2 = 4$$

There are two ways to solve this recurrence. First, without annihilators.

$$r_n = r_{n-1} + n = r_{n-2} + (n-1) + n = r_0 + \sum_{k=1}^{n} k = 1 + \frac{n(n+1)}{2} = 1 + \frac{n+n^2}{2}$$

We can also solve it using the annihilators. The annihilator for the homogeneous part is $E-1$ and for the non-homogeneous part of n it is $(E-1)^2$. So $(E-1)^3$ annihilates the whole sequence. Therefore the recurrence has a form of

$$r_n = c_0 + c_1 n + c_2 n^2$$

Now using the initial conditions

$$\begin{aligned} r_0 &= 1 = c_0 && (1) \\ r_1 &= 2 = c_0 + c_1 + c_2 \Rightarrow c_1 + c_2 = 1 && (2) \\ r_2 &= 4 = c_0 + 2c_1 + 4c_2 && (3) \end{aligned}$$

Subtracting (2) from (3) twice we get

$$2c_2 = 1 \Rightarrow c_2 = 1/2$$

Then from (2) we get

$$c_1 = 1/2$$

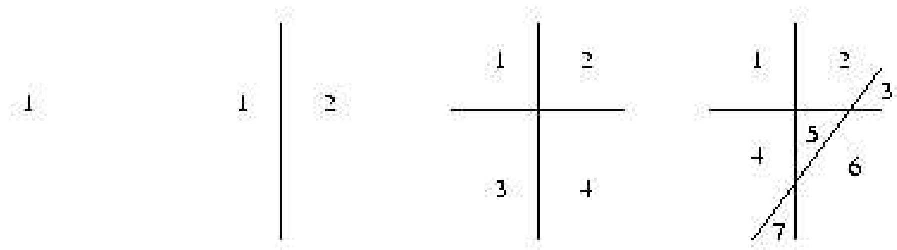
Thus the exact solution is

$$r_n = 1 + 1/2n + 1/2n^2 = 1 + \frac{n+n^2}{2}$$

Common mistakes:

- Just writing the summation and not the closed form for the first method
- Arithmetical errors for the annihilator method

(d) Let c_n be the number of *closed* regions into which n intersecting lines cut the plane. What are c_0, c_1, c_2, c_3?
Solution: 1 point You need at least three lines to make a closed region, so $c_0 = c_1 = c_2 = 0, c_3 = 1$



With three lines, region 5 is closed.
Common mistakes:

- Considering the plane a closed region
- Drawing parallel lines

(e) Obtain a recurrence relation for c_n. Justify.
Solution: 3 points Since you need at lest three lines to make a closed region, as we draw the new nth line, only the regions between the intersection points with the old lines are closed. With $n-1$ old lines there are $n-1$ intersection points and therefore $n-2$ regions between them. The regions before the first intersection point and after the last intersection point are not closed since they are formed by only two lines. So there are $n-2$ new closed regions in addition to the old c_{n-1} ones. Thus the recurrence is:

$$c_n = c_{n-1} + n - 2$$

Common mistakes:

- Justifying the recurrence by "I see the pattern in the sequence..."
- Justifying the recurrence by "A new line adds $n-2$ new regions" without saying why there are $n-2$ new regions
- "The new line cuts every old closed region in half and adds one more closed region so the recurrence is $c_n = 2c_{n-1} + 1$"

(f) Solve the recurrence for c_n to obtain a formula for the number of closed regions. You may use the answer from part (c).

Solution: Again, there are several ways to solve the recurrence. First, we can convert it to the same type of summation as in part (c):

$$c_n = c_{n-1} + n - 2 = c_{n-2} + (n-3) + (n-2) = c_2 + \sum_{k=1}^{n-2} k = \frac{(n-2)(n-1)}{2} = \frac{2 - 3n + n^2}{2}$$

We can solve the recurrence using annihilators. In fact, since the homogeneous part is the same as in (c) and the non-homogeneous part is of the same degree, it has the same annihilator $(E-1)^3$ and the same form

$$c_n = t_0 + t_1 n + t_2 n^2$$

However, we cannot use the first two base cases to find the constants (if you tried, you probably find that out by getting zeroes). This is because you need three lines to form a closed region, so the first equation with the closed region is $c_3 = c_2 + 1$. Thus c_2 is the first base case we can use. We will need three base cases to find the three constants, so we need to calculate $c_4 = c_3 + 2 = 3$.

$$\begin{aligned} c_2 &= 0 = t_0 + 2t_1 + 4t_2 && (4) \\ c_3 &= 1 = t_0 + 3t_1 + 9t_2 && (5) \\ c_4 &= 3 = t_0 + 4t_1 + 16t_2 && (6) \\ & && (7) \end{aligned}$$

Subtracting (1) from (2) we get

$$t_1 + 5t_2 = 1 \Rightarrow t_1 = 1 - 5t_2$$

Subtracting (1) from (3) twice we get

$$-t_0 + 8t_2 = 3 \Rightarrow t_0 = 8t_2 - 3$$

Substituting both of these into (1) we get

$$2t_2 = 1 \Rightarrow t_2 = 1/2 \Rightarrow 1_0 = 1, t_1 = -3/2$$

Therefore

$$c_n = 1 - 3/2n + 1/2n^2 = \frac{2 - 3n + n^2}{2}$$

We can solve the recurrence in yet another way by noticing that

$$c_n = r_{n-2} - 1$$

Therefore, using the solution for r_n from (c)

$$c_n = 1 + \frac{(n-2) + (n-2)^2}{2} - 1 = \frac{n - 2 + n^2 - 2n + 4}{2} = \frac{2 - 3n + n^2}{2}$$

Common mistakes:

- Just writing the summation and not the closed form for the first method
- Arithmetical errors for the annihilator method
- Using the wrong relationship between c_n and r_n for the third method