La Soirée Groupée

Année 2021 – 2022

Cyril BOULLIS, Alexis FIGUERAS, élèves de classe 2°01 et de classe 2°08.

Encadrés par Hélène COCHARD et Cécile DAMONGEOT pour MATh.en.JEANS

Établissement : Lycée Blaise Pascal Orsay

Chercheur : Lucas ERTZBISCHOFF, École polytechnique, Palaiseau.

1. Présentation du sujet

Dans une soirée, il y a $n \epsilon N^*$ invités à l'instant n. On pose $a>0$ donné et fixé. A l'instant suivant, un nouvel invité arrive à la fête. Des groupes se créent dès le début de la soirée selon la règle suivante : à l'instant $n+1$, le nouvel invité arrive parmi les n personnes déjà présentes et décide :
- soit de créer un nouveau groupe et de le rejoindre, avec la probabilité $\frac{a}{a+n}$;
- soit de rejoindre un groupe déjà existant et comportant k invités (où $k \epsilon N^*$), avec la probabilité $\frac{k}{a+n}$

Une fois qu'un invité s'installe dans un groupe, il ne le quitte plus au cours de la soirée. De plus, aucun invité ne quitte la soirée. On laisse alors la soirée évoluer selon la règle précédente.

À l'instant n, y a-t-il beaucoup de groupes composés d'une seule personne, ou au contraire, observe-t-on souvent un unique groupe d'invités ? Comment se répartissent les groupes en général à l'instant n ? Et si on laisse la soirée se dérouler ?

2. Annonce des conjectures et résultats obtenus

Conjecture :
Si $a<1$, on observe souvent peu de groupes avec beaucoup d'invités, surtout si la valeur de a est très faible.
Si $a \approx 1$, on observe souvent un groupe important avec de nombreuses personnes et de petits groupes à côté de peu de personnes
Si $a>1$, on observe souvent de nombreux petits groupes de peu de personnes, surtout si la valeur de a est très grande.

Nous avons créé un programme Scratch qui permet de visualiser l'évolution du nombre de groupes et leurs tailles, et qui nous a permis d'émettre les conjectures ci-dessus.

Nous avons également obtenu une formule, grâce à notre chercheur qui nous l'a donnée car nous ne la trouvions pas. Cependant, nous allons la démonter à la fin de cet article. Cette formule permet, pour une répartition donnée d'un certain nombre d'invités, de calculer la probabilité d'obtenir cette répartition. Elle nous a été très utile dans la résolution du problème.

3. Texte de l'article

1. Des probabilités dépendantes de a

Nous allons dans cette partie montrer que le choix de a est primordial pour les répartitions finales.

1. a, une valeur qui change tout

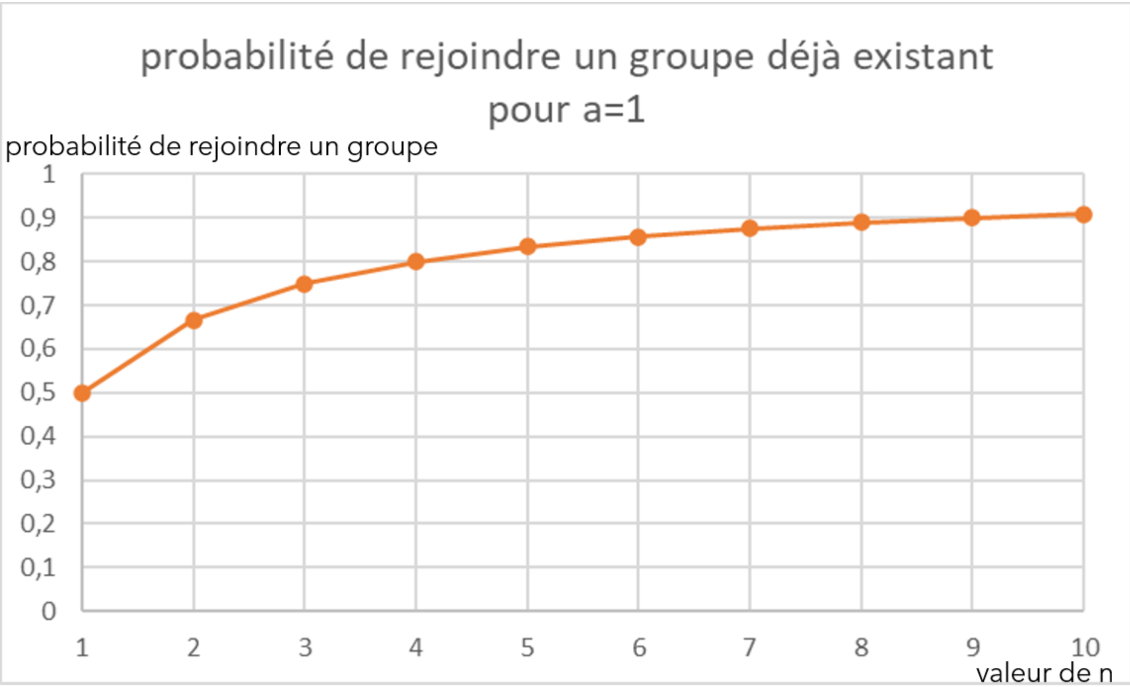
Nous avons très vite remarqué que a définissait très fortement les probabilités. À l'instant $n+1$:

- $\frac{a}{a+n}$ Probabilité de créer un nouveau groupe.
- $1-\frac{a}{a+n}$:Probabilité de rejoindre un groupe (quel qu'il soit).

Une valeur de a très basse donnera donc une probabilité très faible de créer un groupe, et une valeur très haute donnera une probabilité très forte de créer un groupe. Pour avoir un cas intéressant, c'est-à-dire un cas où on risque peu d'obtenir une répartition avec un seul unique gros groupe ou plein de groupes d'une seule personne, le choix de la valeur de a est donc très important.(1)

2. Le cas le plus simple : $a=1$

Lorsque $a=1$, les probabilités sont relativement simples à calculer et ont peu de chances de donner une répartition avec un seul unique gros groupe ou plein de groupes d'une seule personne. C'est la meilleure valeur pour commencer.

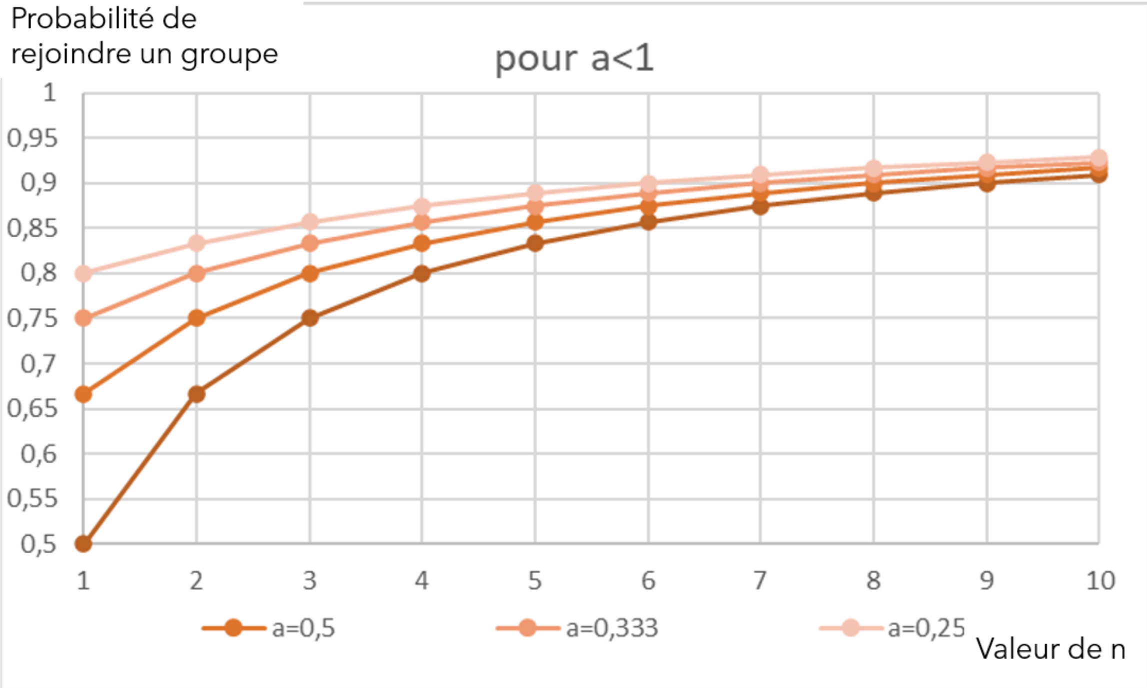


On observe que la probabilité de rejoindre un groupe augmente au fur et à mesure de l'augmentation de n, et tend vers 1. Nous pouvons même le démontrer :

Nous avons tracé les points d'ordonnées $1-\frac{1}{1+n}$. $1+n$ est croissant, donc $\frac{1}{1+n}$ est décroissant et tend vers 0, et donc logiquement $1-\frac{1}{1+n}$ est croissant et tend vers 1.

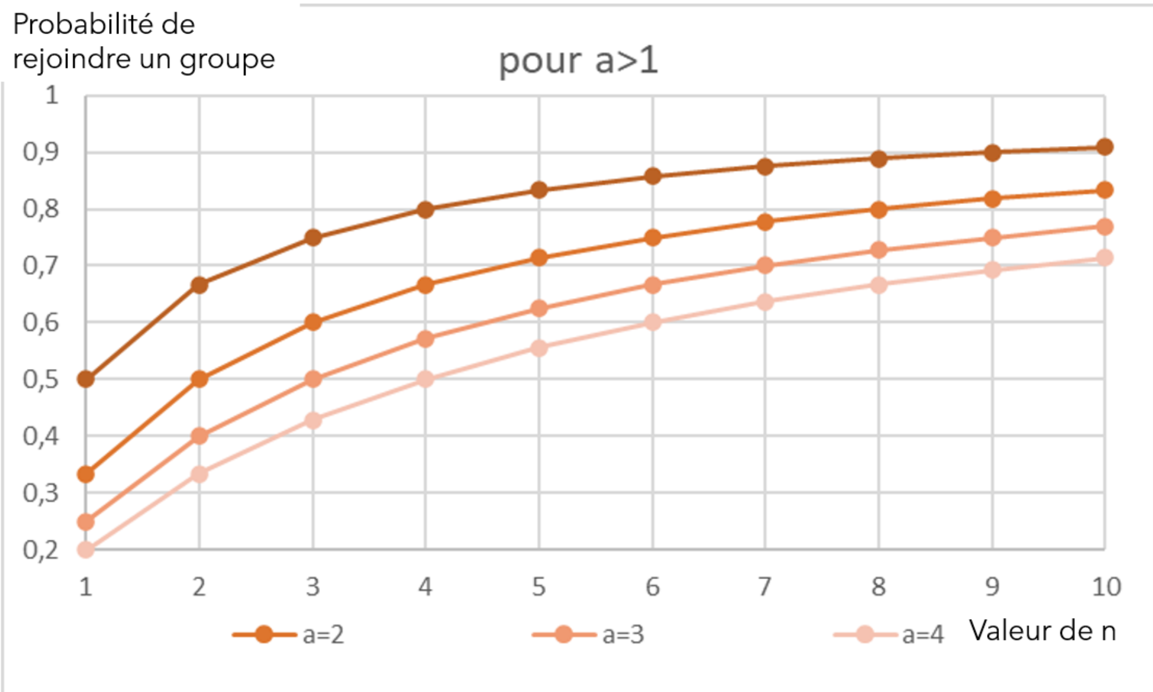
3. Les cas où $a<1$

Nous avons laissé dans les deux graphiques suivants la courbe du cas $a=1$, afin de pouvoir comparer. C'est dans les deux cas la courbe la plus foncée.



On observe que la probabilité de rejoindre un groupe augmente au fur et à mesure de l'augmentation de n, et tend vers 1. Plus a est petit, plus la probabilité de créer un groupe est faible, donc celle de rejoindre sera au contraire plus forte. Pour des valeurs basses de n, la différence est très visible, mais les courbes se rejoignent lorsque n augmente.

4. Les cas où $a>1$



On observe que la probabilité de rejoindre un groupe augmente au fur et à mesure de l'augmentation de n, et tend vers 1, comme dans les cas où $a<1$. Mais au contraire, plus a est grand plus la probabilité de créer un groupe est haute et donc celle de rejoindre un groupe est basse. À la différence du cas $a<1$, on observe une forte différence entre les courbes des différentes valeurs de a : les courbes des valeurs plus hautes de a sont bien en-dessous des autres, ce qui montre une probabilité bien plus faible de rejoindre un groupe.

2. Les partitions(2)

Dans cette partie, nous allons définir et utiliser une notion : les partitions. Elles vont nous permettre de représenter des répartitions possibles des invités à la soirée, et de définir plus tard d'autres notions qui vont nous servir pour la formule.

1. Définitions

On nomme X[n] l'ensemble des nombres entiers de 1 jusqu'à n.

Voici la définition de ce qu'est une partition de X[n] :

- Une partition G de X[n] est un ensemble de parties non vides de X[n] deux à deux disjointes et dont l'union est X[n].

Pour simplifier, chaque élément de la partition représente un invité et chaque groupe représente un groupe d'invités.

Par exemple voici différentes partitions de X[5] :
{{1 ; 4}{2}{3 ; 5}}, {{1 ; 2}{3 ; 4 ; 5}}, {{1 ; 2 ; 3 ; 4 ; 5}}, {{1}{2}{3}{4}{5}}
Dans le premier cas, on a la première personne arrivée dans un groupe avec la quatrième, la deuxième personne est seule, et les personnes 3 et 5 sont ensemble.

On nommera G_n la partition aléatoire en groupes à l'instant n, et G la partition donnée dont on calcule la probabilité d'être égale à G_n.

On utilise la notation |G| pour désigner le nombre de groupes de la partition G, et |g| pour désigner le nombre de personnes dans le groupe g.

2. La partition G'

On nomme G' une partition de l'ensemble X[$n+1$] qui peut être obtenue à partir de la partition G de X[n] donnée, suivant la règle d'ajout de {n+1} dans un groupe déjà existant ou dans son groupe seul.

Si G de X[5] = {{1 ; 4}{2}{3 ; 5}}, alors G' de X[6] peut être égal à :
{{1 ; 4 ; 6}{2}{3 ; 5}}, {{1 ; 4}{2 ; 6}{3 ; 5}}, {{1 ; 4}{2}{3 ; 5}{6}}
Mais G' ne peut pas être égal à :
{{1 ; 6}{2 ; 4}{3 ; 5}}, car on a changé l'élément {4} de groupe alors que c'est interdit.(3)

3. Le passage de G à G'

La probabilité de passer de G à G' vaut :

- $\frac{a}{a+n}$ si la partition G' peut être obtenue à partir de la partition G en ajoutant l'élément {n+1} seul à G.
  Dans ce cas, G' = G U {n+1}.
- $\frac{h}{a+n}$ si la partition G' peut être obtenue à partir de la partition G en ajoutant l'élément {n+1} dans le groupe h de G. Dans ce cas, G' = (G \ h) U (h U {n+1})

3. La formule

1. Présentation de la formule

N'arrivant pas à trouver de moyen de calculer la probabilité d'obtenir une répartition donnée, notre chercheur nous a donné cette formule qui permet de calculer cette probabilité en fonction de a :

$\mathrm{P}(G_n = \mathrm{G}) = \dfrac{a^{|G|}}{a(a+1)(a+2)\ldots(a+n-1)} \prod_{g \in G} (|g|-1)!$

où :

- $|G|$ est le nombre de groupes de G
- g est un groupe de G
- $g \in G$ signifie qu'on l'applique pour tous les groupes de G
- $|g|$ est le nombre de personnes dans le groupe g
- ! est la fonction factorielle : on prend le nombre devant et on le multiplie par tous les entiers précédents jusqu'à 1 inclus (par exemple 4 ! = $1 \times 2 \times 3 \times 4$). On ne peut appliquer la factorielle qu'à un entier naturel, mais cela ne pose pas de problème pour notre formule car $|g|$ est un entier naturel non nul, et donc $|g|-1$ est un entier naturel.

2. Exemple

Pour la partition $\mathrm{G} = \{\{1\ ;\ 2\}\{3\}\}$:

$$\begin{aligned}
\mathrm{P}(G_3 = \mathrm{G}) \quad &= \frac{a^{|G|}}{a(a+1)(a+2)\ldots(a+n-1)} \prod_{g \in G} (|g|-1) \\
&= \frac{a^2}{a(a+1)(a+2)} \times (|g1|-1)! \times (|g2|-1)! \\
&= \frac{a^2}{a(a+1)(a+2)} \times (2-1)! \times (1-1)! \\
&= \frac{a^2}{a(a+1)(a+2)} \times 1! \times 0! \\
&= \frac{a^2}{a(a+1)(a+2)}
\end{aligned}$$

Or si on calcule avec les probabilités qui ont été données au départ on a :

- La personne 1 qui crée son groupe, on multiplie par $\frac{a}{a+0}$.
- La personne 2 qui rejoint le groupe composé d'une seule personne, on multiplie par $\frac{1}{a+1}$.
- La personne 3 qui crée son groupe, on multiplie par $\frac{a}{a+2}$.

On obtient finalement $\frac{a}{a+0} \times \frac{1}{a+1} \times \frac{a}{a+2} = \frac{a^2}{a(a+1)(a+2)}$, ce qui est le même résultat que celui trouvé via la

¿

formule.

3. Démonstration

Nous allons démontrer la formule, en utilisant la démonstration par récurrence. La démonstration par récurrence est une démonstration dans laquelle on démontre que quelque chose est vrai au rang 0 ou 1 selon le cas, puis que si c'est vrai au rang n alors c'est vrai au rang $n+1$. On démontre donc que c'est vrai pour tous les rangs. Nous allons utiliser la notation $g1$ pour désigner le premier groupe de G, $g2$ pour le second, etc.

Démontrons tout d'abord qu'elle est vraie au rang $n=1$: la seule partition existante de X[1] est G = {{1}}. La probabilité d'obtenir cette répartition est normalement 1, car G est la seule possibilité de partition. Calculons la probabilité donnée par la formule :

P(G1 = G) $= \frac{a^{|G|}}{a(a+1)(a+2)\dots(a+n-1)} \prod_{g\in G} (|g|-1)!$

$= \frac{a^1}{a} \times (|g1|-1)!$

$= 1 \times (1-1)!$

$= 0!$

$= 1$

La formule est donc vraie au rang 1.

Nous allons ensuite démontrer que, si elle est vraie au rang n, alors elle est vraie au rang $n+1$.
Nous supposons donc qu'elle est vraie au rang n.
Nous allons également faire une disjonction de cas, en séparant le cas où la personne $n+1$ crée un groupe et celui où elle rejoint un groupe.

Dans le cas où l'élément $\{n+1\}$ rejoint la partition G pour donner la partition G' en créant un groupe, alors on a $|\{n+1\}| = 1$, $G' = G \cup \{n+1\}$ et $|G'| = |G| + 1$, et on multiplie par $\frac{a}{a+n}$, qui représente la probabilité de passer de G à G' dans le cas où $\{n+1\}$ crée son propre groupe, comme nous l'avons vu dans la section 2. 3.

P(Gn+1 = G') $= \frac{a^{|G|}}{a(a+1)(a+2)\dots(a+n-1)} \prod_{g\in G} (|g|-1)! \times \frac{h}{a+n}$

$= \frac{a^{|G|+1}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G} (|g|-1)! \times (|h'|-1)!$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G, g\notin h\text{¿}} (|g|-1)! \times (|h|-1)! \times (|h'|-1)$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G, g\notin h} (|g|-1)! \times (|h'|-2)! \times (|h'|-1)$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G, g\notin h} (|g|-1)! \times (|h|-1)! \times (|h'|-1)$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G'} (|g|-1)!$

On a donc bien la formule au rang $n+1$.

Dans le cas où l'élément $\{n+1\}$ rejoint la partition G pour donner la partition G' en rejoignant un groupe h, alors on a $h' = h \cup \{n+1\}$, $|h'| = |h| + 1$, $G' = G\backslash h \cup h'$ et $|G'| = |G|$, et on multiplie par $\frac{|h|}{a+n}$, qui représente la probabilité de passer de G à G' dans le cas où {n + 1} rejoint un groupe, comme nous l'avons vu dans la section 2. 3.

P(Gn+1 = G') $= \frac{a^{|G|}}{a(a+1)(a+2)\dots(a+n-1)} \prod_{g\in G} (|g|-1)! \times \frac{|h|}{a+n}$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G} (|g|-1)! \times (|h'|-1)$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G, \notin h} (|g|-1)! \times (|h|-1)! \times (|h'|-1)$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G, \notin h} (|g|-1)! \times (|h'|-2)! \times (|h'|-1)$

$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g\in G, \notin h} (|g|-1)! \times (|h'|-1)!$

$$= \frac{a^{|G'|}}{a(a+1)(a+2)\dots(a+n-1)(a+n)} \prod_{g \in G'} (|g|-1)! \qquad =$$

On a donc bien la formule au rang $n+1$.

On obtient donc que si la formule est vraie au rang n, alors elle est vraie n $+¿1$.

On a donc la formule qui est vraie au rang 1, et si elle est vraie au rang n, alors elle est vraie $n+1$. La formule est donc vraie pour tous les n ¿ N*.

4. Conclusion

Dans cette soirée, les tailles et nombre de groupes se définissent principalement par a : avec une valeur faible (par exemple 0,1) créer un groupe est peu probable alors qu'avec une valeur haute (par exemple 10) créer un groupe est très probable. Mais la probabilité de créer un groupe diminue au cours du temps, même avec une valeur élevée.

La formule nous a été très utile afin de comparer les probabilités de certaines répartitions en fonction de a. Grâce à cette formule, nous avons pu émettre des conjectures intéressantes :

- Si a est petit, on a très souvent un très gros groupe et quelques petits groupes à côté.
- Si a est grand, on a très souvent plein de petits groupes de 1 ou 2 personnes.
- Plus on laisse la soirée continuer, plus on a souvent 1 groupe qui grossit et monopolise presque tous les nouveaux arrivants, et ce groupe est souvent le premier, parfois le deuxième, groupe créé. Cela arrive même lorsque a est élevé, car la probabilité de rejoindre un groupe diminue avec le temps.

Nous n'avons pas démontré ces conjectures mais nous avons créé un programme Scratch qui modélise notre sujet, et qui permet de les visualiser.(4)

Notes d'édition

(1) La définition aurait pu être plus claire et contenir une vérification que la somme des probabilités vaut bien 1.

(2) Cette phrase est un peu confuse. Si on veut rester précis avec la terminologie mathématique, une partition ayant été définie comme un ensemble de parties, les éléments d'une partition sont les groupes d'invités et ce sont les entiers de 1 à n qui représentent les invités.

(3) Il n'est pas dit que toutes les possibilités sont données, mais comme il n'y en a pas beaucoup il serait plus clair d'être exhaustif : il y a aussi la possibilité {{1 ; 4}{2}{3 ; 5 ; 6}}.

(4) Il aurait pu être intéressant de noter que la probabilité ne dépend que du nombre de groupes et de leurs tailles, mais pas de la répartition exacte des entiers dans ces groupes. On pourrait associer à une partition des entiers de 1 à n au sens de l'article une partition de n au sens de la combinatoire, comme façon d'écrire n en somme d'entiers non-nuls et où l'ordre ne compte pas. Ce sont les seules informations dont dépendent la probabilité. Par exemple, on remarque que {{1 ; 4}{2}{3 ; 5 ; 6}} a la même probabilité que {{1}{2 ; 6}{3 ; 4 ; 5}}, parce que dans les deux cas on partitionne {1;2;3;4;5;6} en un ensemble de taille 3, un ensemble de taille 2 et un ensemble de taille 1, ce que l'on résume par 6 = 3 + 2 + 1.