MS-A409 Grundkurs i diskret matematik
Appendix, del I

G. Gripenberg

Aalto-universitetet

2 oktober 2013

Varför är inte mängdläran så enkel som den ser ut?

Om man endast behandlar mängder av typen $\{1, 3, 4, 7\}$ med ändligt många element blir det inga större problem och de mängdteoretiska beteckningarna kan vara behändiga att använda. Det enklaste och klassiska exemplet som visar att det kan bli problem är att försöka definiera en mängd A med

$$A = \{\, x : x \notin x \,\}.$$

Om $A \in A$ så gäller inte villkoret $A \notin A$ och enligt defintionen av A gäller då $A \notin A$ och man får en motsägelse. Om däremot $A \notin A$ så uppfyller A villkoret fö mängden A och man har $A \in A$ vilket igen är en motsägelse. Så hur man än vänder sig får man en motsägelse och det är inte bra! En vardaglig motsvarighet är påståendet "Detta är en lögn" eller att tala om "Barberaren som klipper håret på alla som inte klipper sitt eget hår".

Vad kan man kräva av ett predikat $L(x, y)$ som säger att "x och y är lika"?

Om $x = y$ så är förstås $y = x$, och om dessutom $x = z$ så är $y = z$. Vidare skulle det vara orimligt att inte ha $x = x$. Om vi nu skriver $x == y$ istället för $L(x, y)$ ser vi att följande satser skall vara sanna:

- $\forall x \forall y((x == y) \to (y == x))$
- $\forall x \forall y \forall z((x == y)\ \&\ (y == z) \to (x == z))$
- $\forall x(x == x)$

Observera att detta är "samma" villkor som för en ekvivalensrelation!

Koordinaterna i ett ordnat par

Den första koordinaten i $[x, y]$ (eller (x, y)) är (förstås) x och den andra y. Om man skriver paret med mängdbeteckningar som $\{\{a\}, \{a, b\}\}$ så kan man definiera predikat $F(p, x)$ och $A(p, y)$ som säger att första koordinaten i p är x och att andra koordinaten i p är y på följande sätt:

$$F(p, x) = \forall z((z \in p) \to (x \in z))$$

(eller kortare $\forall z \in p(x \in z)$) och

$$A(p, y) = (\exists z((z \in p)\ \&\ (y \in p)))\ \&$$
$$(\forall u(\forall v(((u \in p)\ \&\ (v \in p))\ \&\ (!(u == v)) \to ((!(y \in u)) \mid (!(y \in v)))))).$$

I detta uttryck finns onödigt många parenteser och man kan också skriva det som $\exists z((z \in p)\ \&\ (y \in p))\ \&\ \forall u \forall v((u \in p)\ \&\ (v \in p)\ \&\ !(u = v) \to !(y \in u) \mid !(y \in v))$. Ännu kortare skulle vara att skriva $\exists z \in p(y \in p)\ \&\ \forall u \in p \forall v \in p(!(u == v) \to (y \notin u) \mid (y \notin v))$. Endel av parenteserna, som tex. kring $y \in u$ kunde man också lämna bort.

Listor, talföljder och kartesiska produkter som funktioner

- *En lista $[a, b, c, d]$ kan tolkas som en funktion f definierad i mängden $\{1, 2, 3, 4\}$ (eller $\{0, 1, 2, 3\}$) så att $f(1) = a$, $f(2) = b$, $f(3) = c$ och $f(4) = d$.*
- *En oändlig talföljd $(a_n)_{n=0}^{\infty} = (a_0, a_1, a_2, \ldots)$ kan tolkas som en funktion f definierad i N_0 så att $f(n) = a_n$ för alla $n \in \mathbb{N}_0$.*
- *Om X_j är en mängd för varje $j \in J$ där J är en (annan) mängd så kan man definiera den kartesiska produkten $\prod_{j \in J} X_j$ som mängden av alla funktioner $f : J \to \bigcup_{j \in J} X_j$ så att $f(j) \in X_j$.*

Hur många jämförelser behövs för att sortera n tal i storleksordning?

Vi skall visa att det räcker med högst $n \log_2(n)$ jämförelser. Då $n = 1$ (eller $n = 2$) är det klart att detta är sant. Antag nu att det stämmer för alla $n \le k$. (Detta är en variant av induktionsprincipen!) och att vi har $k + 1$ tal som vi skall ordna. Dela upp dessa i två mängder med m tal som vi ordnar och sedan kombinerar vi dessa ordnade mängder till en mängd. De senare kan göras med $2m - 1$ jämförelser och genom att använda induktionsaxiomet får vi att det sammanlagda antalet jämförelser blir högst

$$2m \log_2(m) + 2m - 1 = 2m \log_2(2m) - 1 + 2m - 1$$
$$= 2m(\log_2(2m) - 1) + 2m - 1 \le 2m \log_2(2m) = (k + 1) \log_2(k + 1).$$

Om $k + 1 = 2m + 1$ så delar vi upp mängden i två mängder med m och $m + 1$ element och får på samma sätt att antalet jämförelser blir högst

Hur många jämförelser behövs för att sortera n tal i storleksordning? Forts.

$$m \log_2(m) + (m + 1) \log_2(m + 1) + 2m$$
$$= m \log_2(m(m + 1)) + \log_2(m + 1) + 2m$$
$$\le m(\log_2(2m + 1)^2) - \log_2(4)) + \log_2(m + 1) + 2m$$
$$\le m(2 \log_2(2m + 1) - 2) + \log_2(2m + 1) + 2m$$
$$\le 2m2 \log_2(2m + 1)) + \log_2(2m + 1) = (k + 1) \log_2(k + 1).$$

Detta innebär att induktionssteget fungerar och att antalet jämförelser som behövs för att ordna n tal hör till mängden $O(n \log_2(n)$ följer av induktionsprincipen.

Hur många jämförelser behövs för att hitta talet med storleksordningsnummer p i en mängd med n tal?

Det är klart att om $p = 1$ (det minsta talet) eller $p = n$ (det största talet) så räcker det med (men behövs också) $n - 1$ jämförelser. Vi skall nu visa att maximimantalet jämförelser då $1 \leq p \leq n$ hör till $O(n)$ dvs. vi skall visa att det finns en konstant C så man genom att göra högst Cn jämförelser kan hitta talet med storeleksordningsnummer p. Om tex. $n \leq 2^{30}$ kan man först sortera talen i storleksordning med högst $30n$ jämförelser, och sedan välja det tal man söker så för dessa fall gäller påståendet med $C = 20$ och vi skall nu visa att detta gäller för alla n. Antag nu att påståendet (med $C = 20$) gäller för alla $n \leq k$ (där $k \geq 2^{20}$). Nu skall vi visa att det också gäller då $n = k + 1$.
Välj m så att $5(2(m-1)+1)+1 \leq k+1 \leq 5(2m+1)$ och lägg till tal som är större alla ursprungliga så vi har sammanlagt $5(2m+1)$ tal. Dela in dessa i $2m+1$ grupper med fem tal och bestäm medianerna av dessa tal. För detta behövs det $6(2m+1)$ jämförelser. Sedan bestämmer vi medianerna av medianerna och kallar detta q. Enligt induktionsantagandet kan detta göras med högst $20(2m+1)$ jämförelser.

Hur många jämförelser behövs för att hitta talet med storleksordningsnummer p i en mängd med n tal?. Forts.

Sedan delar vi in alla de övriga talen i två mängder, dels de som är större än q och de som är mindre än q. Vi vet att $3m+2$ av talen är större än q och $3m+2$ är mindre så vi behöver bara $5(2m+1)-1-6m-4 = 4m$ jämförelser för detta. Dessutom, och detta är viktigare, den större av dessa mängder innehåller högst $5(2m+1)-1-3m-2 = 7m+2$ tal. Nu är det tal vi söker antingen q eller ligger i någondera av dessa mängder så vi skall i värsta fall bestämma ett tal med ett visst ordningsnummer i en mängd med $7m+2$ element. Enligt induktionsantagandet kan detta göras med högst $20(7m+2)$ jämförelser. Nu har vi alltså visat att vi kan hitta det tal vi sökt med högst $6(2m+1)+20(2m+1)+10m+20(7m+2)$ jämförelser. För att induktionssteget skall fungera måste detta tal vara mindre än $20(k+1)$ och eftersom $k+1 \geq 5(2(m-1)+1)+1$ så räcker det att visa att

$$6(2m+1)+20(2m+1)+4m+20(7m+2) \leq 20(5(2(m-1)+1)+1).$$

Denhär olikheten gäller då $m \geq \frac{146}{4}$ vilket säkert är fallet då $2m+1 \geq 2^{20}$.

Ordnat val av r föremål från en mängd med n föremål

- *Om varje föremål kan väljas bara en gång kan det första väljas på n olika sätt, det andra på $n-1$ olika sätt och så vidare så att föremål nummer r kan väljas på $n-r+1$ olika sätt. Genom att använda produktregeln ser vi att antalet olika möjligheter är possibilities is $n \cdot (n-1) \cdot (n-2) \cdot \ldots (n-r+1)$ eller $\dfrac{n!}{(n-r)!}$.*
- *Om varje föremål kan väljas flera gånger (dvs. de tas inte bort ur mängden eller så är mängdens element typer av föremål som tas från något annat ställe) då finns det n alternativ vid varje val så att det följer av produktregeln att antalet möjligheter är n^r.*

Icke-ordnat val av r föremål från en mängd med n föremål **utan** upprepningar

Låt $b(n,r)$ vara detta antal av icke-ordnade val av r föremål från en mängd med n föremål utan upprepningar. När vi har gjort ett sådant val får vi ett ordnat val genom att ordna de valda r föremålen. Detta kan göras på $r!$ olika sätt så det följer av produktprincipen att antalet sätt göra att ett ordnat val av r föremål från en mängd med n föremål utan upprepningar är $b(n,r)cdotr!$. Eftersom vi vet att detta antal är $n \cdot (n-1) \cdot \ldots \cdot (n-r+1) = \frac{n!}{(n-r)!}$ så får vi

$$b(n,r) = \frac{n!}{r! \cdot (n-r)!} = \binom{n}{r}.$$

Icke-ordnat val av r föremål från en mängd med n föremål **med** upprepningar, I

Antag att mängden är $\{a_1, a_2, \ldots, a_n\}$. Om vi har en lista med de $r+n-1$ talen $1, 2, \ldots, r+n-1$ kan vi välja välja $n-1$ av dessa och sätta dem i växande ordning $1 \leq p_1 < p_2 < \ldots < p_{n-1} \leq r+n-1$. Låt nu $r_1 = p_1 - 1$, $r_2 = p_2 - p_1 - 1$, …, $r_k = p_k - p_{k-1} - 1$, …, $r_n = r + n - 1 - p_{n-1}$. Det ger nu ett val där vi valt r_k föremål av typ a_k och $\sum_{k=1}^{n} r_k = r$. Å andra sidan, om vi väljer r_k föremål av typ a_k så att $\sum_{k=1}^{n} r_k = r$ då bestämmer detta ett val av talen p_j så att $p_1 = r_1 + 1$, och $p_k = p_{k-1} + r_k + 1$, $k = 2, \ldots, n-1$. Därför kommer antalet icke-ordnade val av r föremål från en mängd med n föremål **med** *upprepningar att bli antalet icke ordnade val av $n-1$ (eller r) föremål från en mängd $r+n-1$ föremål* **utan** *upprepningar, dvs.* $\binom{r+n-1}{n-1} = \binom{r+n-1}{r}$.

Icke-ordnat val av r föremål från en mängd med n föremål **med** upprepningar, II

Låt $f(r, n)$ vara antalet sätt på vilka man kan göra ett icke-ordnat val av r föremål från en mängd med n föremål med upprepningar. Ett sådant val är detsamma som att placera r föremål i n ordnade (dvs. inte identiska) lådor. Om $n = 1$, så kan detta göras på bara ett sätt så att $f(r, 1) = 1$ för alla $r \geq 0$. Om $n > 1$ kan vi sätta $j = 0, 1, \ldots, r$ föremål i den första lådan och de återstående $r - j$ föremålen i de återstående $n-1$ lådorna. Eftersom vi får olika resultat för varje val värde på j så får vi rekursionsekvationen

$$f(r, n) = \sum_{j=0}^{r} f(r-j, n-1) \overset{k=r-j}{=} \sum_{k=0}^{r} f(k, n-1).$$

I synnerhet betyder detta att $f(r, 2) = r + 1$ och med hjälp av formeln för summan av en aritmetisk serie får vi $f(r, 3) = \frac{(r+2)(r+1)}{2}$. Nu gissar vi att $f(r, n) = \binom{r+n-1}{n-1} = \binom{r+n-1}{r}$ så vi skall visa att

$$\binom{r+n-1}{n-1} = \sum_{k=0}^{r} \binom{k+n-2}{n-2}, \quad r \geq 0, \quad n \geq 2.$$

Icke-ordnat val av r föremål från en mängd med n föremål **med** upprepningar, II, forts.

Denna likhet gäller säkert för $r = 0$ och varje $n \geq 2$. Antag att den gäller för $r = s$ och $n \geq 2$. Då får vi när $r = s + 1$ och $n \geq 2$

$$\begin{aligned}
\sum_{k=0}^{s+1} \binom{k+n-2}{n-2} &= \binom{s+1+n-2}{n-2} + \sum_{k=0}^{s} \binom{k+n-2}{n-2} \\
&= \binom{s+1+n-2}{n-2} + \binom{s+n-1}{n-1} \\
&= \frac{(s+n-1)\cdot\ldots\cdot(s+2)}{(n-2)!} + \frac{(s+n-1)\cdot\ldots\cdot(s+1)}{(n-1)!} \\
&= \frac{(s+n-1)\cdot\ldots\cdot(s+2)\cdot(n-1+s+1)}{(n-1)!} \\
&= \frac{(s+n)\cdot(s+n-1)\cdot\ldots\cdot(s+n-(n-1)+1)}{(n-1)!} = \binom{s+1+n-1}{n-1}.
\end{aligned}$$

Induktionssteget fungerar och påståendet följer med induktionsprincipen.

Antalet surjektioner $A \to B$ då $|A| = m$ och $|B| = n$

Antag att $B = \{y_1, y_2, \ldots, y_m\}$. Låt $F = B^A$ vara mängden av alla funktioner $A \to B$. Låt $F_j = (B \setminus \{y_j\})^A \subset F$ vara mängden av alla funktioner $A :\to B \setminus \{y_j\}$, dvs. alla funktioner $f \in F$ så att $f(x) \neq y_j$ för alla $x \in A$. Detta innebär att mängden av surjektioner är $F \setminus \cup_{j=1}^{m} F_j$. Nu är $F_{j_1} \cap F_{j_2} \cap \ldots \cap F_{j_k}$ mängden $(B \setminus \{y_{j_1}, y_{j_2}, \ldots y_{j_k}\})^A$ av alla funktioner $A \to B$ som inte får något av värdena $y_{j_1}, \ldots, y_{j_k}$. Om $1 \leq j_1 < \ldots < j_k \leq n$ så är $|F_{j_1} \cap F_{j_2} \cap \ldots \cap F_{j_k}| = (n-k)^m$. Eftersom indexen $1 \leq j_1 < \ldots < j_k \leq n$ kan väljas på $\binom{n}{k}$ olika sätt så kan vi med hjälp av inklusions-exklusionsprincipen dra slutsatsen att antalet surjektioner: $A \to B$ är

$$n^m - \left(\sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)^m\right) = \sum_{r=0}^{n} (-1)^{n-r} \binom{n}{r} r^m.$$

Observera att då $m < n$ så finns det inga surjektioner: $A \to B$ så att $\sum_{r=0}^{n} (-1)^{n-r} \binom{n}{r} r^m = 0$ då $n < m$, vilket kanske inte är helt uppenbart.