Złożoność komunikacyjna

Karolina Taborek

21 grudnia 2004 roku

$Relacja\ R \subseteq X \times Y \times Z$

$Alicja\ zna\ tylko\ x \in X$
$Bob\ zna\ tylko\ y \in Y$

Zadanie

$obliczyć\ z \in Z\ takie\ ,\ że\ (x\ ,\ y\ ,\ z) \in R$

Definicja

Protokół P oblicza relację R, jeżeli dla wszystkich poprawnych danych wejściowych $(x, y) \in X \times Y$ protokół osiąga liść etykietowany wartością $z \in Z$ taką, że $(x, y, z) \in R$.

Definicja

$D(R)$ – deterministyczna złożoność komunikacyjna relacji R: liczba bitów przesłanych przez najlepszy protokół obliczający relację R w przypadku najgorszych danych wejściowych (legalnych bądź nie).

Kosztem protokołu jest wysokość drzewa.

Definicja

$A \times B$ *jest monochromatycznym prostokątem*
$(w\ odniesieniu\ do\ relacji\ R \subseteq X \times Y \times Z)$,
jeśli istnieje taka wartość $z \in Z$,
że dla każdej pary $(x, y) \subseteq A \times B$
albo $(x, y, z) \in R$ *albo* (x, y) *jest niepoprawna.*

	000	001	010	011	100	101	110	111
000	∅	{3}	{2}	{2,3}	{1}	{1,3}	{1,2}	{1,2,3}
001	{3}	∅	{2,3}	{2}	{1,3}	{1}	{1,2,3}	{1,2}
010	{2}	{2,3}	∅	{3}	{1,2}	{1,2,3}	{1}	{1,3}
011	{2,3}	{2}	{3}	∅	{1,2, 3}	{1,2}	{1,3}	{1}
100	{1}	{1,3}	{1,2}	{1,2,3}	∅	{3}	{2}	{2,3}
101	{1,3}	{1}	{1,2,3}	{1,2}	{3}	∅	{2,3}	{2}
110	{1,2}	{1,2,3}	{1}	{1,3}	{2}	{2,3}	∅	{3}
111	{1,2, 3}	{1,2}	{1,3}	{1}	{2,3}	{2}	{3}	∅

Przykładowy podział

Twierdzenie

Każdy $t-$bitowy protokół P obliczający relację R indukuje podział zbioru $X \times Y$ na co najwyżej 2^t monochromatycznych prostokątów.

Przykład

$$R=\{(x,y,m):|x\cap y|-n/12 \leqslant m \leqslant |x\cap y|+n/12,\ x,y\subseteq\{1,2..,n\}\}$$

Pokażemy , że $D(R)=\Omega(n)$
poprzez wykazanie , że istnieje
zbiór oszukujący rozmiaru $t=2^{(\Omega(n))}$

Rozwiązanie:

$S_1, S_2, \ldots, S_t$ – *losowe podzbiory zbioru* $\{1, 2, \ldots, n\}$

Rozpatrujemy pary wejść :

$(S_1, \bar{S}_1), (S_2, \bar{S}_2), \ldots, (S_t, \bar{S}_t)$

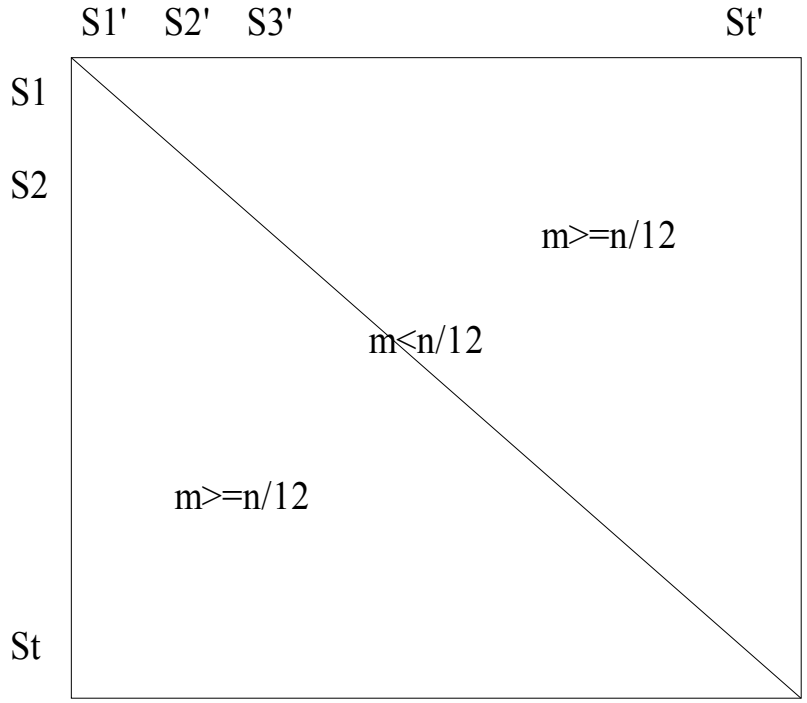
Przy takim wyborze zbiorów , że :

$|S_i \cap \bar{S}_j| > n/6 \;\; \forall\, i \neq j$

żadne z tych par nie mogą być w tym samym monochromatycznym prostokącie.

$$R=\{(x,y,m):|x\cap y|-n/12 \leqslant m \leqslant |x\cap y|+n/12,\ x,y\subseteq\{1,2..,n\}\}$$

$$(S_i,\bar{S}_i,m)\in R \Rightarrow m\leqslant n/12,\ \textit{ponieważ}\ |S_i\cap\bar{S}_i|=0$$
$$(S_i,\bar{S}_j,m)\in R, i\neq j \Rightarrow m>n/12,\ \textit{ponieważ}\ |S_i\cap\bar{S}_j|>n/6\ \textit{dla}\ i\neq j$$



Taki wybór istnieje.

$S_i, S_j - losowe\ podzbiory\ zbioru \{1,2\ldots, n\}$

$Z_k - zmienna\ losowa\ o\ wartości\ 1,\ gdy\ k \in S_i\ i\ k \in S_j\ i\ 0\ wpp.$

$$E[Z_k] = 1 * Pr(Z_k = 1) + 0 * Pr(Z_k = 0) = Pr(k \in S_i \wedge k \in S_j) = 1/4$$

Korzystając z nierówności Chernoffa otrzymujemy :

$$Pr[\sum_{k=1}^{n} Z_k \leqslant n/6] \leqslant Pr[\left|(\sum_{k=1}^{n} Z_k)/n - 1/4\right| \geqslant 1/12] \leqslant 1/2^{cn}$$

dla pewnej stałej c.

$Pr[\left|S_i \cap S_j\right| \leqslant n/6] \leqslant 1/2^{cn}$ *dla pewnej stałej c.*

$$Pr\left(\left|S_i \cap \bar{S}_j\right| \leqslant n/6\right) \leqslant t^2 * 2^{-cn} \text{ dla pewnych } i, j$$

$$\text{dla } t = 2^{(cn/2)} \text{ mamy :}$$

$$Pr\left(\left|S_i \cap \bar{S}_j\right| \leqslant n/6\right) < 1 \text{ dla pewnych } i, j$$
$$\text{zatem } Pr\left(\left|S_i \cap \bar{S}_j\right| > n/6\right) > 0 \ \forall\, i, j$$

$$\text{stąd istnieje } t = 2^{(\Omega(n))} \text{ zbiorów } S_i \text{ takich , że pary}$$
$$(S_1, \bar{S}_1), (S_2, \bar{S}_2), \ldots, (S_n, \bar{S}_n) \text{ muszą należeć}$$
$$\text{do różnych prostokątów monochromatycznych.}$$

Przykład

Relacja uniwersalna

$$U=\{(x,y,i): x\in\{0,1\}^n, y\in\{0,1\}^n, i\in\{1,2,\ldots,n\}, x_i\neq y_i\}$$

pary (x,x) *są niepoprawnymi danymi wejściowymi*

$$D(U)\leqslant n+\log n$$

Alicja wysyła x do Boba. Bob znajduje odpowiedni indeks.

$$D(U)\geqslant D(NE)-2=n-1,$$

gdzie $NE(x,y)=[x\neq y]$

Zakładając , że mamy protokół P_U obliczający U
konstruujemy protokół dla NE w następujący sposób :

– *Alicja i Bob korzystając z protokołu* P_U
dla danych wejściowych (x, y)

– *jeżeli* P_U *nie zwróci wyniku* $i \in \{1, 2, \ldots, n\}$,
wtedy wynikiem jest 0

– *jeżeli* P_U *zwróci jako wynik* $i \in \{1, 2, \ldots, n\}$,
wtedy Alicja przesyła do Boba x_i
Bob zwraca 1 *jeżeli* $x_i \neq y_i$, 0 *wpp.*

$-\ x \neq y$

protokół P_U *gwarantuje zwrócenie i takiego , że* $x_i \neq y_i$,
zatem wynikiem konstruowanego protokołu będzie 1

$-\ x = y$

protokół P_U *może zwrócić* 0 *lub* $i \in \{1, 2, \ldots, n\}$
wynikiem konstrukowanego protokołu będzie 0

$$D(U) \geqslant D(NE) - 2 = n - 1$$

Przykład

$$X = \{ x_1 x_2 \ldots x_n : \forall i \; x_i \in \{0, 1\ \}, \; (\sum_{i=1}^{n} x_i) \bmod 2 = 0 \}$$

$$Y = \{ y_1 y_2 \ldots y_n : \forall i \; y_i \in \{0, 1\ \}, \; (\sum_{i=1}^{n} y_i) \bmod 2 = 1 \}$$

Relacja parzystości :

$$P = \{ (x, y, i) : x \in X, y \in Y, i \in \{1,2 \ldots, n\}, x_i \neq y_i \}$$

takie indeks zawsze istnieje , ponieważ $x \neq y \; (X \cap Y = \emptyset)$

Pokażemy , że $D(P) \leqslant 2 \log n$

Alicja i Bob binarnie wyszukują i takiego , że $x_i \neq y_i$

W każdym kroku :

- Alicja i Bob mają przedział $[\, j \,,\, k \,]$ taki ,
  że parzystości $x_j \ldots x_k$ i $y_j \ldots y_k$ są różne
- $l := (j+k)/2$
- Alicja wysyła do Boba parzytość $x_j \ldots x_l (1 \ bit)$
  Bob wysyła do Alicji parzystość $y_1 \ldots y_l (1 \ bit)$
- jeżeli parzystości są różne ,
  to ustawiają $k = l$ i kontynuują
- jeżeli parzystości są równe , to wnioskują ,
  że parzystości $x_{l+1} \ldots x_k$ i $y_{l+1} \ldots y_k$ są różne
  więc ustawiają $j = l+1$ i kontynuują

P wykonuje $\log n$ kroków.
W każdym kroku przesyłane są 2 bity
i kiedy zbiór jest jednoelementowy ,
indeks $j = k$ jest szukanym indeksem.

Jest to najlepszy możliwy protokół dla P.

Lemat

Niech X , Y *będą rozłącznymi podzbiorami* $\{0,1\}^n$,
$C=\{(x,y): x\in X, y\in Y, d(x,y)=1\}$,
gdzie $d(x,y)$ *– odległość Hamminga między* x *i* y
i $R=\{(x,y,i): x\in X, y\in Y, x_i\neq y_i\}$

wtedy
liczba podziałów R *spełnia nierówność :*
$C^D(R)\geqslant(|C|)^2 / |X|*|Y|$

Dowód

$R_1, R_2, \ldots, R_t - monochromatyczne\ prostokąty$
$(dla\ relacji\ R)\ optymalnego\ podziału\ X \times Y$
$m_i - ilość\ C - elementów\ należących\ do\ R_i$
$|R_i| - ilość\ elementów\ w\ prostokącie\ R_i$

Zachodzą równości :

$$|C| = \sum_{i=1}^{t} |m_i|$$

$$\sum_{i=1}^{t} |R_i| = |X| * |Y|$$

$$|R_i| \geqslant m_i^2$$

W każdym rzędzie x i w każdej kolumnie y prostokąta R_i znajduje się co najwyżej jeden $C-$element.

$j-$*wynik odpowiadający prostokątowi* R_i

Wszystkie $y-$*ki w prostokącie* R_i *różnią się od* x *na* $j-$*tym bicie.*
Jeżeli $(x, y)\in C$, *to* $d(x, y)=1$,
zatem x *i* y *różnią się na dokładnie jednym bicie,* $j-$*tym bicie.*

$$(|C|)^2 = (\sum_{i=1}^{t} m_i)^2$$

$$(\sum_{i=1}^{t} m_i)^2 \leqslant t * \sum_{i=1}^{t} m_i^2 \ (nierówność\ Cauchy - Schwartz)$$

$$t * \sum_{i=1}^{t} m_i^2 \leqslant t * \sum_{i=1}^{t} |R_i|$$

$$t * \sum_{i=1}^{t} |R_i| = t * |X| * |Y|$$

Otrzymujemy :

$$t \geqslant (|C|)^2 / |X| * |Y|$$

Relacja parzystości :

$$P=\{(x,y,i): x\in X, y\in Y, i\in\{1,2,\ldots,n\}, x_i\neq y_i\}$$

spełnia założenia powyższego lematu.

Zachodzą równości :

$$|X|=|Y|=2^{(n-1)}$$

$$|C|=n*2^{(n-1)} \quad (\forall x\in X\, d(x,y)=1 \quad \Rightarrow y\in Y)$$

Na mocy lematu otrzymujemy :

$$C^D(R)\geqslant(|C|)^2/|X|*|Y|$$

$$C^D(R)\geqslant n^2*(2^{(n-1)})^2/(2^{(n-1)})^2=n^2$$

stąd $D(R)\geqslant 2\log n$

Przykład

Relacja uniwersalna :

$$U=\{(x,y,i): x\in\{0,1\}^n, y\in\{0,1\}^n, i\in\{1,2,\ldots,n\}, x_i\neq y_i\}$$

$$R(U)=O(\log n)$$

$Alicja\ i\ Bob\ podejmują\ t\ prób:$

$-\ wybierają\ losowy\ ciąg\ r \in \{0,1\}^n$

$-\ Alicja\ wysyła\ do\ Boba\ iloczyn\ skalarny < x, r > \ (1\ bit)$

$-\ Bob\ wysyła\ do\ Alicji\ iloczyn\ skalarny < y, r > \ (1\ bit)$

$-\ jeżeli < x, r > \neq < y, r >, to\ ograniczają\ swoje\ ciągi$
$do\ tych\ bitów\ i, na\ których\ r_i = 1$

$Parzystości\ otrzymanych\ ciągów\ sa\ różne.$
$Korzystają\ z\ (deterministycznego)\ wyszukiwania\ binarnego$
$do\ znaleznienia\ i\ takiego, że\ x_i \neq y_i.$

$Protokół\ zwraca\ jako\ wynik\ indeks\ i.$

$-\ jeżeli\ w\ t\ próbach\ ponoszą\ porażkę\ przy\ wyborze$
$ciągu\ r\ takiego, że < x, r > \neq < y, r >, to\ zwracają$
$dowolny\ indeks\ i$

$Liczba\ wymienionych\ bitów\ wynosi\ 2t + O(\log n)$

Prawdopodobieństwo błędu wynosi 2^{-t},
ponieważ dla $x \neq y$ *i dla losowego ciągu* $r \in \{0, 1\}^n$

$$Pr(\langle x, r\rangle \neq \langle y, r\rangle) = Pr(\langle |x-y|, r\rangle \neq 0) = 1/2$$

Dla $t = \log n$ *otrzymujemy złożoność* $O(\log n)$
z prawdopodobieństwem błędu $1/n$.

Przykład

Relacja monotoniczna :

$$M = \{(x, y, i) : x \in \{0,1\}^n, y \in \{0,1\}^n, i \in \{1, 2, \ldots, n\}, x_i = 1, \ y_i = 0 \ \}$$

(pary , dla których takie i nie istnieje są nielegalne)

$$D(M) \leqslant n + \log n$$

analogicznie do relacji U

$$D(M) \geqslant D(DISJ) - 2 = n - 1,$$

gdzie DISJ jest funkcją rozłączności

Zakładając , że mamy protokół P_M *obliczający M kontruujemy protokół dla DISJ w następujący sposób :*

- *Bob konstruuje ciąg* y' *spełniający warunek :*
  $\forall i\ y'_i = 1 - y_i \left(\exists i\ x_i = y_i = 1 \iff \exists i\ x_i = 1,\ y'_i = 0\right)$
- *Alicja i Bob korzystając z protokołu* P_M *dla* (x, y') *otrzymując wynik i*
- *Alicja przesyła do Boba* x_i
- *jeżeli faktycznie* $x_i = 1\ i\ y'_i = 0,$ *to Bob zwraca* $0,$ *wpp. zwraca* 1

$-x \cap y \neq \emptyset$

protokół P_M *gwrantuje zwrócenie indeksu i takiego ,*
że $x_i = 1$ *i* $y'_i = 0$, *więc wynikiem będzie* 0

$-x \cap y = \emptyset$

niezależnie od wyniku zwróconego przez P_M
wynikiem będzie 1, *ponieważ Alicja przesyła* x_i
do Boba , który dokonuje sprawdzenia

$$D(DISJ) \leqslant D(M) + 2$$
$$D(M) \geqslant D(DISJ) - 2 = n - 1$$

Przykład

Relacja parami rozłączne :

$n = 3m$

$X = \{P : P = \{(x_1, y_1), \ldots, (x_m, y_m) : \forall i\, x_i, y_i \in \{1,2,\ldots,n\}, \forall i\, x_i \neq y_i, \forall i, j, i \neq j\, x_i \neq x_j\}\}$

$Y = \{S : S \subseteq \{1,2,\ldots,n\}, |S| = m-1\}$

$PR = \{(P, S, i) : P \in X, S \in Y, i-ta\ para\ P\ nie\ ma\ elementu\ ze\ zbioru\ S\}$

$(definicje\ zbiorów\ P\ i\ S\ zapewniają\ istnienie\ takiego\ i)$

$D(M) = \Omega(m)$

$$M' = \{(P, S, i) \in M : \forall (x_k, y_k) \in P \ \neg(x_k \in S \wedge y_k \in S)\}$$

M ' otrzymaliśmy z M dzięki ograniczeniu danych wejściowych , zatem jest to prostszy problem , stąd :

$$D(M') \leqslant D(M)$$

Rozważmy następujący problem:

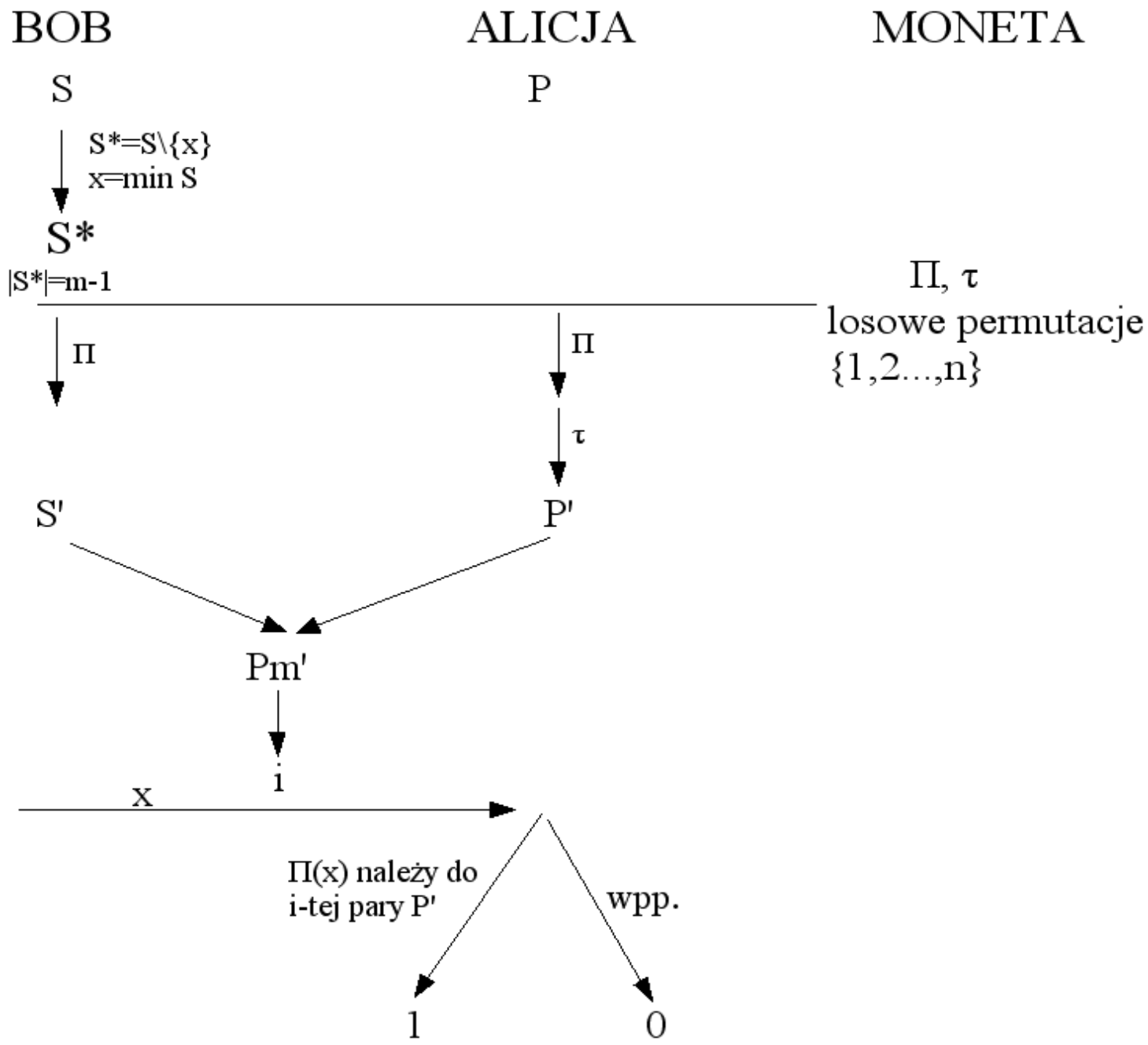
Bob zna zbiór S *(tym razem rozmiaru* m*)*
Alicja zna zbiór P
określone jak w relacji M'*, tzn. żadna para*
z P *nie zawiera dwóch elementów ze zbioru* S

Zadanie:
Obliczyć funkcję częściową f *taką, że:*
– $f(P, S) = 0$
jeżeli w P *jest para, która nie ma elementu z* S
– $f(P, S) = 1$
jeżeli każda para w P *ma dokładnie jeden element z* S

Lemat

$$R_{1/4}^{pub}(f) \leqslant 2(D(M') + \log n)$$

Zakładając, że mamy dany deterministyczny protokół $P_{M'}$ *obliczający* M', *konstruujemy randomizowany protokół* P_f, *który oblicza* f *ze złożonością* $D(M') + \log n$ *i który popełnia błąd z prawdopodobieństwem co najwyżej* $1/2$ (*popełnia błąd tylko wtedy, gdy wynikiem jest* 0). *Dwukrotne wykonanie redukuje prawdopodobieństwo błędu do* $1/4$.


BOB
ALICJA
MONETA
S
P
S*=S\{x}
x=min S
S*
|S*|=m-1
Π, τ
losowe permutacje
{1,2...,n}
Π
Π
τ
S'
P'
Pm'
i
x
Π(x) należy do
i-tej pary P'
wpp.
1
0

– $f(P, S) = 1$

tzn. każda para P ma dokładnie jeden element z S. Po usunięciu x ze zbioru S jest dokładnie jedna para p_i' w P', która nie ma elementu ze zbioru S' oraz $\Pi(x) \in p_i'$.

Wynikiem protokołu P_M, będzie dokładnie ten indeks i, zatem w tym przypadku Alicja zawsze poda poprawną odpowiedź 1.

$-\ f(P, S) = 0$

$tzn.\ istnieje\ co\ najmiej\ jedna\ para\ w\ P,$
$ozn.\ p_k, która\ nie\ ma\ elementu\ z\ S$

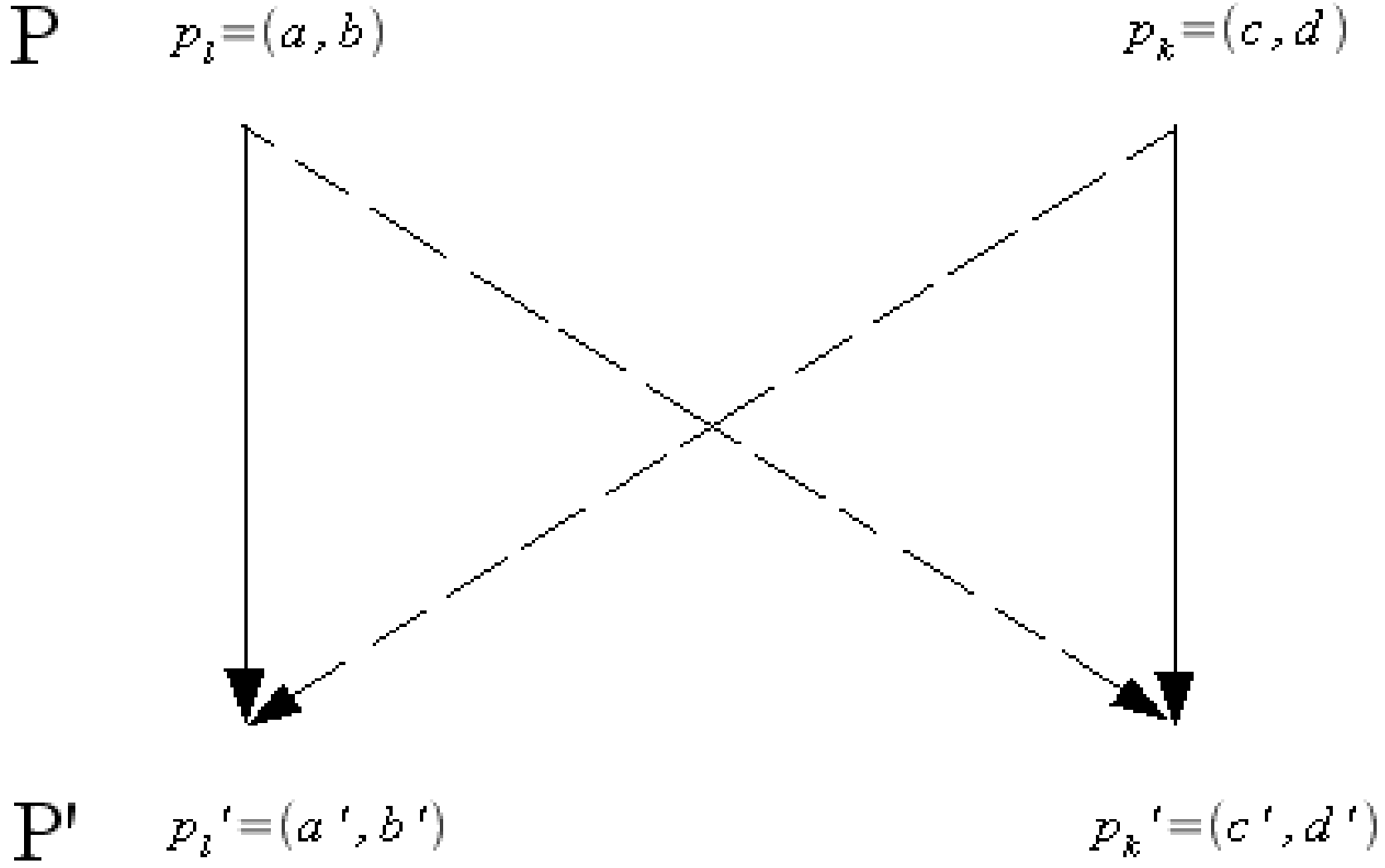
$-\ x\ nie\ jest\ elementem\ żadnej\ pary\ z\ P,$
$zatem\ \Pi(x)\ nie\ jest\ elementem\ żadnej$
$pary\ P',\ więc\ \Pi(x)\ nie\ jest\ elementem$
$i-tej\ pary\ znalezionej\ przez\ protokół\ P_{M'}$
$Alicja\ zawsze\ poda\ poprawną\ odpowiedź\ 0$

$-\ x\ jest\ elementem\ p_l \in P$
$w\ tym\ przypadku\ protokól\ P_{M'}\ podaje$
$błędny\ wynik\ z\ prawdopodobieństwem$
$co\ najwyżej\ 1/2$

x jest elementem pary $p_l \in P$

W tym przypadku istnieją co najmniej dwie
pary p_k, $p_l \in P'$, które nie mają elementu S'.
Protokół P_M, z prawdopodobieństwem co najmniej $1/2$
znajduje parę, która nie zawiera $\Pi(x)$, ponieważ
istnieje różnowartościowe przekształcenie,
które dla każdej pary permutacji Π, τ przyporzadkowuje
parę Π', τ' taką, że otrzymane zbiory (P', S') są takie same,
ale $\Pi(x)$ i $\Pi'(x)$ należą do różnych par P'.

Permutacje są losowe, zatem niezależnie od wyniku i
otrzymanego dla (P', S') z prawdopodobieństwem co najmniej
$1/2$ element $\Pi(x)$ nie należy do $i-$tej pary P'.


P
$p_i=(a,b)$
$p_k=(c,d)$
P'
$p_i'=(a',b')$
$p_k'=(c',d')$

Lemat

$R^{pub}_{1/4}(DISJ) \leqslant R^{pub}_{1/4}(f),$
gdzie DISJ jest funkcją rozłączności
dla danych wejściowych z $\{0,1\}^m \times \{0,1\}^m$

Zakładając , że mamy protokół P_f *obliczający f*
skonstruujemy protokół dla DISJ mający taką samą
złożoność i prawdopodobieństwo błędu.

- *Alicja znając* $x \in \{0,1\}^m$ *konstruuje zbiór*
  $P = \{(3i - x_i - 1, 3i) : 1 \leqslant i \leqslant m\}$
- *Bob znając* $y \in \{0,1\}^m$ *konstruuje zbiór*
  $S = \{3i - y_i : 1 \leqslant i \leqslant m\}$
- *Alicja i Bob korzystają z protokołu* P_f
  dla danych wejściowych (P, S)

- $DISJ(x, y) = 0$

  $x \cap y \neq \emptyset \Rightarrow \exists i \; x_i = y_i = 1$

  $p_i = (3i - 2, 3i), \; s_i = 3i - 1$

  $\forall i \neq j \; s_j \notin p_i$

  zatem p_i *nie ma elementu z* S,

  P_f *zwróci w tym przypadku* 0

- $DISJ(x, y) = 1$

  $x \cap y = \emptyset \Rightarrow \forall i \; (y_i = 0) \vee (y_i = 1 \wedge x_i = 0)$

  $\forall i \; (s_i = 3i) \vee (s_i = 3i - 1 \wedge p_i = (3i - 1, 3i))$

  $\forall i \; s_i \in p_i$

  każda para ma element z S,

  P_f *zwróci wartość* 1

Prawdopodobieństwa sukcesu protokołów
P_f *i protokołu dla DISJ są takie same.*

$$R^{pub}_{1/4} = \Omega(m)$$

$$D(M) \geq D(M')$$

$$D(M') = \Omega(R^{pub}_{1/4}(f) - logm) = \Omega(R^{pub}_{1/4}(DISJ) - logm) = \Omega(m)$$

$$D(M) = \Omega(m)$$