Bollettino
Unione Matematica Italiana

Sezione A – La Matematica nella Società e nella Cultura

Fabio Burderi

Sulla decifrabilità dei codici

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. **9-A**—*La Matematica nella Società e nella Cultura* (2006), n.2 (Fascicolo dedicato alle tesi di dottorato), p. 219–222.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2006_8_9A_2_219_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2006.

Bollettino U. M. I.
La Matematica nella Società e nella Cultura
Serie VIII, Vol. IX-A, Agosto 2006, 219-222

Sulla decifrabilità dei codici

Fabio Burderi

1. – Introduzione.

La parola *codice* usualmente denota in letteratura (cf [1]) un insieme di parole su un fissato alfabeto tale che ogni messaggio, cioè una concatenazione di tali parole, ha un'unica decodifica in termini di parole di codice; cioè si può risalire in modo unico alla sequenza di parole che costituiscono il messaggio. In questa tesi consideriamo una condizione più generale di decifrabilità e quindi la parola *codice* qui denota solamente un insieme di parole su un fissato alfabeto; per riferirci invece al significato originale, parleremo di codici univocamente decifrabili (*UD*).

Questa tesi consta di due capitoli che corrispondono a due differenti approcci allo studio della nozione di decifrabilità più debole della univoca decifrabilità. Nel primo capitolo studiamo la relazione tra la nozione di varietà di codici, introdotta da Guzmán (cf [4]), e la somma di Kraft di un codice. Ciò porta a caratterizzare la varietà dei codici *UD* in termini di somma di Kraft.

In questo studio introduciamo la nozione di massimalità in una varietà ed estendiamo, ad una varietà generica, risultati riguardanti massimalità e completezza dei codici *UD*.

Nel secondo capitolo introduciamo, per un qualunque codice, la nozione di *coding partition*.

Questa nozione fornisce uno strumento per studiare le ambiguità di un codice. Infatti, dato un codice, possiamo ottenere una sua partizione in classi tali che le ambiguità del codice sono localizzate all'interno di ciascuna classe della partizione e quindi, a livallo delle classi c'è una sorta di univoca decifrabilità: è possibile fattorizzare univocamente un messaggio in blocchi, dove ciascun blocco è una concatenazione di parole appartenenti a una stessa classe della partizione e blocchi consecutivi sono composti da parole appartenenti a classi della partizione differenti tra loro.

Alla fine del capitolo mostriamo un collegamento tra i due approcci, cioè tra le varietà di codici e la coding partition. Noi proviamo che se tutte le classi di una coding partition sono costituite da codici appartenenti ad una stessa varietà, allora l'intero codice appartiene a questa varietà.

2. – Decifrabilità dei codici e disuguaglianza di Kraft.

Lo studio su condizioni di decifrabilità più deboli della univoca decifrabilità (*UD*) è stato introdotto in [5] da Lempel, che ha introdotto la nozione di codice *multiset*

decifrabile (*MSD*). In questo caso l'informazione che interessa è data dal multiset delle parole di codice che formano il messaggio, per cui è consentito che un messaggio abbia più decodifiche a patto che queste condividano tutte lo stesso multiset di parole di codice. In [4] Guzmán considera inoltre la nozione di codice *set decifrabile* (*SD*) secondo cui tutte le decodifiche di un messaggio devono avere lo stesso set di parole di codice. Introduce quindi un concetto molto generale di decifrabilità usando i monoidi e da questo definisce poi le *varietà* di codici. Le classi $\mathcal{UD}$, $\mathcal{MSD}$, $\mathcal{SD}$ risultano casi particolari di *varietà*.

Sia C un codice e sia M un monoide. Diciamo che C è *decifrabile* in M se ogni applicazione $f : C \to M$ si estende a un (unico) omomorfismo $\overline{f} : C^* \to M$.

Data $\mathcal{N}$ classe di monoidi, denotiamo con $\mathcal{C}(\mathcal{N})$ la classe di codici C che sono decifrabili in ogni $M \in \mathcal{N}$. Viceversa data $\mathcal{K}$ classe di codici, denotiamo con $\mathcal{M}(\mathcal{K})$ la classe di monoidi M tale che ogni $C \in \mathcal{K}$ è decifrabile in M.

DEFINIZIONE 1. – *Una classe $\mathcal{V}$ di codici è una varietà di codici se $\mathcal{V} = \mathcal{C}(\mathcal{M}(\mathcal{V}))$.*

La terminologia è motivata dal fatto che $\mathcal{V}$ è una varietà di codici se e solo se $\mathcal{V} = \mathcal{C}(\mathcal{W})$ per qualche $\mathcal{W}$ varietà di monoidi.

Un codice C su un alfabeto finito A è *completo* se ogni parola di A^* è fattore di qualche elemento di C^*; un codice appartenente ad una varietà $\mathcal{V}$, è detto un $\mathcal{V}$-codice; un $\mathcal{V}$-codice è massimale se non è sott'insieme proprio di nessun altro $\mathcal{V}$-codice.

TEOREMA 1. – *Sia $\mathcal{V}$ una varietà di codici e sia C un codice nella varietà $\mathcal{V}$. Se C è un $\mathcal{V}$-codice massimale, allora C è completo.*

TEOREMA 2. – *Sia $\mathcal{V}$ una varietà di codici inclusa nella varietà $\mathcal{SD}$, i.e. $\mathcal{V} \subseteq \mathcal{SD}$, e sia $C \in \mathcal{V}$ un codice regolare. Se C è completo, allora C è un $\mathcal{V}$-codice massimale.*

COROLLARIO 1. – *Siano $\mathcal{V}_1$ e $\mathcal{V}_2$ varietà di codici tali che $\mathcal{V}_1 \subseteq \mathcal{V}_2 \subseteq \mathcal{SD}$, e sia $C \in \mathcal{V}_1$ un codice regolare. Se C è massimale nella vatietà $\mathcal{V}_1$, allora è massimale nella varietà $\mathcal{V}_2$.*

TEOREMA 3. – *Sia $\mathcal{V}$ una varietà di codici e sia $C \in \mathcal{V}$ un codice regolare. Allora esiste un codice regolare e completo Y tale che $C \subseteq Y$ e $Y \in \mathcal{V}$.*

DEFINIZIONE 2. – *Dato un codice C su un alfabeto finito A, con $card(A) = d$, la somma di Kraft di C è data da*

$$K(C) = \sum_{c \in C} d^{-|c|}$$

dove $|c|$ è la lunghezza della parola di codice c.

TEOREMA 4. – *Sia $\mathcal{V}$ una varietà di codici. Se $K(C) \leq 1$ per ogni $C \in \mathcal{V}$, allora $\mathcal{V} = \mathcal{UD}$.*

TEOREMA 5. – *Sia $\mathcal{V}$ una varietà di codici tale che $\mathcal{V} \neq \mathcal{UD}$. Allora, per ogni $\varepsilon > 0$, esiste un codice completo e regolare $Y \in \mathcal{V}$ tale che*

$$1 < K(Y) < 1 + \varepsilon.$$

3. – Coding partition.

Sia X un codice e sia

$$P = \{X_1, X_2, \ldots\},$$

una partizione di X cioè : $\bigcup_{i \geq 1} X_i = X$ e $X_i \cap X_j = \emptyset$, per $i \neq j$.

Diciamo che la partizione P è *concatenatively independent* se, per $i \neq j$,

$$X_i^+ \cap X_j^+ = \emptyset.$$

Sia $P = \{X_1, X_2, \ldots\}$ una partizione concatenatively independent di un codice X. Una *P-fattorizzazione* di un elemento $z \in X^+$ è una fattorizzazione $z = z_1 z_2 \cdots z_t$, dove

- $\forall i \;\; z_i \in X_k^+$, per qualche $k \geq 1$
- if $t > 1$, $z_i \in X_k^+ \Rightarrow z_{i+1} \notin X_k^+$, per ogni $1 \leq i \leq t-1$.

DEFINIZIONE 3. – *Una partizione P è una coding partition se è concatenatively independent e inoltre ogni elemento $z \in X^+$ ha un'unica P-fattorizzazione, cioè se*

$$z = x_1 x_2 \cdots x_s = y_1 y_2 \cdots y_t,$$

dove $x_1 x_2 \ldots x_s$, $y_1 y_2 \ldots y_t$ sono P-fattorizzazioni di z, allora $s = t$ e $x_i = y_i$ per $i = 1, \ldots, s$.

Ricordiamo che è possibile definire il seguente ordine tra le partizioni di un insieme X: se P_1 e P_2 sono due partizioni di X, $P_1 \leq P_2$ se gli elementi di P_1 sono unioni di elementi di P_2.

TEOREMA 6. – *Dato un codice X, esiste la più fine coding partition P di X. Questa è chiamata la partizione caratteristica di X e si denota con $P(X)$.*

Sia X un codice e sia $P(X)$ la partizione caratteristica di X. Sia X_0 l'unione di quelle classi di $P(X)$ formate da un solo elemento. Da $P(X)$ possiamo allora derivare un'altra partizione di X

$$P_C(X) = \{X_0, X_1, \ldots\},$$

dove $|X_i| > 1$, per $i \geq 1$. La partizione $P_C(X)$ è detta la *partizione canonica* di X.

TEOREMA 7. – *Esiste un algoritmo che fornisce la partizione canonica di un codice finito X.*

TEOREMA 8. – *Data una partizione $P = \{X_1, X_2, \ldots, X_n\}$ tale che X_i, per $i = 1, 2, \ldots, n$, è un insieme razionale, allora è decidibile se P è una coding partition.*

CONGETTURA: *Se X è razionale, il numero delle classi di $P_C(X)$ è finito e ciascuna classe di $P_C(X)$ è un insieme razionale.*

TEOREMA 9. – *Sia $P = \{X_1, X_2, \ldots, X_n\}$ una coding partition di un codice X e sia $\mathcal{K}$ una varietà di codici. Se $X_i \in \mathcal{K}$, per $i = 1, 2, \ldots, n$, allora $X \in \mathcal{K}$.*

BIBLIOGRAFIA

[1] J. BERSTEL e D. PERRIN, *The Theory of Codes*, Academic Press, New York (1985).
[2] S. BURRIS e H.P. SANKAPPANAVAR, *A Course in Universal Algebra*, Springer, New York (1981).
[3] S. EILENBERG, *Automata, Languages and Machines, Vol. A*, Academic Press, New York (1974).
[4] F. GUZMÁN, *Decipherability of codes*, Journal of Pure and Applied Algebra, **141** (1999), 13-35.
[5] A. LEMPEL, *On multiset decipherable codes*, IEEE Trans. Inform. Theory, **32** (1986), 714-716.

Dipartimento di Matematica Università di Palermo
e-mail: burderi@math.unipa.it
Dottorato di Ricerca in Matematica (sede amministrativa: Palermo) – Ciclo XV
Direttori di Ricerca: Prof. Antonio Giambruno, Università di Palermo
Prof. Antonio Restivo, Università di Palermo