1

Capitolo C65
informazioni, probabilità discrete, codici

Contenuti delle sezioni

a. informazioni, indagini, probabilità p. 2
b. probabilità discreta p. 6
c. nozione di codice p. 11

14 pagine

C65 0.01 I termini che compaiono nel titolo di questo capitolo ricorrono spesso sui mezzi di comunicazione e nei discorsi quotidiani.
Sono termini che riguardano temi fondamentali, ma che vengono usati in varie accezioni (questo accade soprattutto per la parola "codice") che spesso sono discordanti, non chiaramente definiti e anche mossi da intenti faziosi.

Qui ci serviremo dei tre termini a partire da definizioni che dovrebbero essere ampiamente condivisibili quando si pongono obiettivi relativamente circoscritti; qui ci occupiamo solo di obiettivi raggiungibili affrontando problemi discreti.

C65 a. informazioni, idagini, probabilità

C65 a.01 Si osserva che il termine **informazione** viene usato con diverse accezioni; per fare chiarezza ciascuna di esse va inquadrata in un suo scenario nel quale compaiono operatori impegnati ad affrontare problemi chiaramente definiti.

Cominciamo con il presentare a livello elementare l'azione dell'informare, ossia il far pervenire dei segnali, da parte di un operatore trasmittente con il ruolo di fonte dell'informazione, a un operatore ricevente che svolge il ruolo di utente dell'informazione.
I segnali trasmessi sono realizzati fisicamente con procedimenti opportuni e sono schematizzati matematicamente come sequenze di segni elementari che gli operatori trasmittente e riceventi sanno riconoscere precisamente e coerentemente e che praticamente possono essere utilizzati dal ricevente per prendere decisioni sopra un problema che sta affrontando.

Il termine informazione viene usato anche per identificare il contenuto della azione dell'informare tramite comunicazione, contenuto che viene determinato facendo riferimento al problema che il ricevente affronta e che il trasmittente in una certa misura condivide.
Il contenuto di una informazione basicamente consiste in sequenze di segni ben distinguibili dai due operatori, mentre sul piano pratico e applicativo può essere collegato a elementi del problema che il ricevitore deve risolvere.
In genere si presume una certa concordanza tra trasmettitore e ricevitore anche nella attribuzione di significati e di finalità ai segni elementari, alle singole sequenze di segni e al loro succedersi in un dato contesto.

C65 a.02 In questo capitolo useremo il termine **configurazioni** o anche il termine **situazioni** per denotare entità che possano essere rappresentate da strutture chiaramente costruibili a partire da oggetti innanzi tutto discreti (numeri, stringhe, grafi, vettori numerici, ...) sottoponibili a elaborazioni facilmente definibili e che possano far parte di modelli di scenari reali o realizzabili (grandezze fisiche scalari o vettoriali, risultati di estrazioni d urne, dati riguardanti comportamenti materiali, sociali o culturali, ...).

Al termine **informazione** attribuiamo un primo significato applicativo che si riferisce alle conoscenze possedute da un operatore OI che conduce una **indagine**.

All'inizio dell'indagine OI possiede informazioni sopra l'insieme S_0 delle configurazioni possibili che egli (o esso) ha ragione di aspettarsi; la definizionedi questo insieme di aspettative iniziali lo supponiamo relativamente semplice.
Nel corso dell'indagine Oi riesce a ridurre l'insieme delle possibili configurazioni a sottoinsiemi di S_0 via via più ridotti (si pensa a una indagine condotta correttamente).
Le definizioni di questi insiemi, che alla fine di una indagine di successo si concludono con una sola configurazione, sono necessariamente sempre più stringenti e in genere sono via via più precise e più elborate nella loro formulazione.
Alla progressiva riduzione delle possibilità viene associato un progressivo aumento della informazione raggiunta dalla indagine.

Gli insiemi di configurazioni attese possono essere pensati sia come interamente delimitati, sia all'opposto come gamme di possibilità concernenti configurazioni delle quali si conoscono poche caratteristiche da considerare essenziali che si intendono conoscere meglio con il proseguire dell'indagine.

Le singole configurazioni sono chiamate spesso **campioni**, termine appropriato quando le possibilità riguardano configurazioni effettivamente disponibili, ma adottato anche per le indagini su possibilità soltanto attese o ipotizzate.

Le più semplici indagini consistono in misurazioni auspicabilmente precise sopra un singolo oggetto ben individuato o valutazioni sopra una collezione di procedimenti den definiti avvenuti o ben prevedibili, come le estrazioni da una urna con possibili contenuti precisamente delimitati.
Altre indagini, all'opposto, si svolgono attraverso la esecuzione di elaborate osservazioni sopra fenomeni descrivibili in termini quantitativi compositi o anche richiedenti valutazioni qualitative.

Lo svolgimento delle indagini più impegnative può essere molto articolato, richiedere strumenti elaborati, attenti preparativi, e quindi consumo di rilevanti risorse materiali e umane.

Qui cominciamo a schematizzare ogni indagine come una semplice sequenza di azioni ben distinte che chiamiamo **passi di indagine**.

Se assumiamo un primo punto di vista insiemistico, una indagine si sviluppa attraverso una sequenza di passi, l'i-esimo dei quqli finalizzato alla individuzione all'interno di un certo presumibile insieme di possibili situazioni S_i di uno dei suoi sottoinsiemi $E_{i,1}, E_{i,2}, ...E_{i,k}$, ... che assumerà il ruolo del successivo insieme di possibilità S_{i+1}.

Le operazioni da attuate nei vari passi delle indagini, riguardano prove che possono far parte ad una ampia tipologia; si pensi ad esempio a tests, sondaggi, misurazioni, fotografie, filmati, carotaggi, monitoraggi ripetitivi, comparazioni, valutazioni qualitative, ...

C65 a.03 Le indagini volte a migliorare le informazioni sopra un insieme di configurazioni possono richiedere passi sia di natura sperimentale (osservazioni con strumenti fisici, chimici, tecnologici, analisi biologiche, economiche, sociologiche, linguistiche, comportamentali, ...) che di natura logico-algoritmica (procedimenti deduttivi, argomentazioni su schemi astratti, elaborazioni basate su modelli quantitativi discreti o continui, ...).

In posizione intermedia si collocano le indagini elaborative condotte con il computer attraverso simulazioni o analisi di insiemi di situazioni molto estesi e/o molto articolati (sistemi complessi, big data, ...), anche servendosi di scelte casuali.

Per queste occorre distinguere tra elaborazioni su configurazioni discrete, per le quali non si pongono problemi di approssimazione e si suppone sia possibile dominare i limiti provenienti dalla implementazione per gli interi trattabili, ed elaborazioni su entità soggette ad approssimazioni (tipicamente numeri, intervalli o plurintervalli razionali da pensare derivati tramite discretizzazione da grandezze idealizzate come continue).

In entrambi i casi si effettuano indagini sperimentali; nel primo caso, però, il livello di affidabilità raggiunto dai dispositivi hardware e dalle piattaforme software consente di ritenere trascurabili le eventualità di malfunzionamenti e di considerare che si tratti di indagini su sistemi formali, anch'esse effettuabili servendosi di strumenti software del genere CAS, `computer algebra systems (we)` o del genere AI.

Le elaborazioni su entità soggette ad approssimazioni (in genere portate avanti con metodi di approssimazione variamente collaudati) presentano invece caratteristiche più vicine a quelle delle indagini sperimentali.

C65 a.04 Sui processi di indagine va detto che, come tutti gli altri procedimenti, richiedono **risorse** e va rilevato, in linea generale, che esse sono riconducibili a tre componenti: tempi di elaborazione, memorie impiegate e attività per l'organizzazione delle elaborazioni.

Notiamo esplicitamente che qui ci occupiamo solo di indagini che permettano di distinguere tra le diverse configurazioni possibili senza modificarle.
Non consideriamo quindi indagini su sistemi fisici quantistici, processi che si servono di misurazioni che modificano le situazioni esaminate in misura che non può essere trascurata.

Più avanti ci interessano solo processi di indagine discreti e quindi i classici esempi dei passi di indagine sono le estrazioni di palline da un'urna, i lanci di monete o dadi, lo scoprimento di carte da gioco, le risposte a quiz con risposte predefinite.

C65 a.05 L'aumento delle informazioni ottenuto con un passo di una indagine va collegato alle caratteristiche su cui si è indagato e all'auspicabile conseguente restrizione dell'insieme delle possibilità da prendere in esame e da sottoporre ai passi d'indagine successivi.

Diciamo che una prova porta a una **effettiva crescita di informazione** quando permette di individuare un sottoinsieme delle possibilità precedentemente consentite che restringe effettivamente le possibilità sulle quali si deve indagare ulteriormente.

Le diverse situazioni che si possono individuare nell'ambito delle possibilità consentite, se fosse possibile ripetere i passi di indagine, cioè le osservazioni, si distinguerebbero per maggiore o minore frequenza di ritrovamento; in genere si usa dire anche che si possono ottenere i diversi risultati con maggiore o minore facilità.

Per trattare quantitativamente queste distinzioni si introducono delle misure che associano "maggiore probabilità" alle situazioni che tendono a riscontrarsi più frequentemente, o in altri termini, a quelle che si ottengono con minore difficoltà.

È ragionevole convenire che un esito di una prova che individua una situazione poco probabile produca un aumento di informazione maggiore di quello che si otterrebbe in seguito ad un esito di prova che individua una situazione più probabile.

È anche ragionevole convenire che i risultati più probabili si possano ottenere con un minore impiego di risorse.

C65 a.06 Per precisare le richieste precedenti affermiamo che ogni passo di indagine ha come scopo la individuazione di un sottoinsieme dell'insieme delle possibilità S.
Questo insieme viene detto anche **spazio dei campioni**. La collezione dei sottoinsiemi di S individuabili mediante un'indagine viene detta collezione degli **eventi**: denotiamola con $\mathcal{E}$; $\mathcal{E} \subseteq \mathfrak{P}(\mathsf{S})$.

Nel caso in cui l'insieme delle possibilità finito spesso si può assumere $\mathcal{E} = \mathfrak{P}(\mathsf{S})$; questo può essere più problematico quando ci si trova di fronte a insiemi di possibilità infiniti.

Per trattare quantitativamente un processo di indagine ci serviamo di una funzione $\mathcal{I}$ chiamata **misura di informazione** che valuta le informazioni fornite dalle prove attuate nei vari passi di una indagine e di una funzione $\mathcal{P}$ detta **misura di probabilità** o anche **legge di probabilità** che valuta le probabilità delle situazioni individuate.

C65 a.07 Vediamo ora quali caratteri generali devono avere queste funzioni di valutazione della probabilità.

Innanzi tutto esse associano agli eventi numeri reali nonnegativi e, per le precedenti considerazioni comparative, si chiede che

$$\forall E_1, E_2 \in \mathcal{E} \quad : \quad E_1 \subset E_2 \implies \mathcal{I}(E_1) > \mathcal{I}(E_2) \ \wedge \ \mathcal{P}(E_1) \leq \mathcal{P}(E_2) \ .$$

I valori assunti da $\mathcal{P}$ devono essere compresi tra 0 ed 1.

Il valore 1 è assunto dall'intero insieme S: questo corrisponde al fatto che si considerano passi di indagine appropriati, cioè effettivamente in grado di individuare solo situazioni facenti parte di S.

Dato che l'indagine fornisce informazioni effettive solo se riesce a distinguere un sottoinsieme proprio di S, si assume che sia $\mathcal{I}(\mathsf{S}) = 0$, cioè che si tratta una quantità di informazione che all'inizio dell'indagine è nulla.

Un $E \subset \mathsf{S}$ è un evento sse si sa attuare una serie di prove che consente di individuare lo stesso E quando le prove ottengono una determinata serie di risposte; la stessa serie di prove, nel caso ottenga una serie di risposte diversa dalla precedente, permette di individuare un evento contenuto nel sottoinsieme complementare $\mathsf{S} \setminus E$.
È lecito e opportuno assumere che anche questo complementare sia un evento, ovvero assumere che sia $\mathsf{S} \setminus E \in \mathcal{E}$.

Se si sanno individuare due sottoinsiemi E_1 ed E_2, si sa individuare anche $E_1 \cap E_2$: questo si può ottenere effettuando le due serie di prove che aprono la possibilità di individuare, risp., E_1 ed E_2, ottenendo proprio le risposte che caratterizzano, risp., E_1 ed E_2.

Osserviamo che questa schematizzazione dipende dall'ipotesi che la effettuazione di ciascuna delle prove non modifica le situazioni che si intendono distinguere.

C65 a.08 Due eventi E_1 ed E_2 si dicono **eventi mutuamente esclusivi** sse corrispondono a due insiemi di campioni disgiunti.
Per essi è lecito chiedere che anche $E_1 \dot\cup E_2$ sia un evento e che sia $\mathcal{P}(E_1 \dot\cup E_2) = \mathcal{P}(E_1) + \mathcal{P}(E_2)$.
Questa richiesta risulta opportuna anche in quanto rende possibile gestire stadi di indagine nei quali sono aperte le due possibilità rappresentate dai due insiemi E_1 ed E_2.

Per ogni evento E accade che E ed $\mathsf{S}\setminus E$ sono mutuamente esclusivi e quindi si assume anche $\mathcal{P}(\mathsf{S}\setminus E) = 1 - \mathcal{P}(E)$.

La collezione $\mathcal{E}$ degli insiemi dotati di probabilità di riscontro dunque contiene anche l'insieme vuoto e deve essere $\mathcal{P}(\emptyset) = 0$; anche questo si cfa dipendere dalla appropriatezza delle indagini: queste devono condurre necessariamente a qualche evento di S che non deve essere necessariamente un singoletto.

Inoltre in genere si assume che $\mathcal{I}$ in corrispondenza di $\emptyset$ assuma il valore improprio $+\infty$.

C65 b. probabilità discreta

C65 b.01 Una indagine discreta riguarda uno spazio dei tipi di campioni S costituito da un insieme finito o numerabile di configurazioni discrete e ha come scopo la determinazione di una singola configurazione o di un insieme di tali oggetti ben circoscritto.

Nella realtà una indagine discreta si può sviluppare attraverso una sequenza finita di prove, ma per trattare più agevolmente i possibili insiemi di indagini conviene ammettere le possibilità di successioni numerabili (potenzialmente infinite) di prove.

Lo svolgimento di una indagine è individuato dalla sequenza limitata o illimitata dei risultati delle prove.
Fissate opportune convenzioni, il risultato di ogni prova si può esprimere mediante un carattere di un alfabeto finito o numerabile, oppure mediante una stringa finita o illimitata facente parte di un determinto linguaggio.
Con questi assunti lo spazio dei campioni può considerarsi un linguaggio e gli eventi suoi sottolinguaggi.

C65 b.02 Nel seguito considereremo prevalentemente processi finiti, cioè procedimenti che effettuano la individuazione di sequenze finite di sottoinsiemi di un dato insieme finito, insieme i cui elementi sono campioni attribuibili a tipi determinati.
I tipi dei campioni si suppone possano essere chiaramente riconosciuti dall'indagatore, mentre per lui i singoli campioni non sono distinguibili; l'indagatore quindi si può trovare di fronte a repliche per ciascuno dei tipi di campioni.
Per una prova effettuata entro uno di questi processi di indagine si usa spesso il termine **estrazione**, termine che rispecchia quello che accade nei processi che consistono in estrazioni di palline da un sacchetto o di bussolotti da un'urna, oggetti appartenenti a tipi distinguibili che possono presentarsi replicati.

Quando saranno trattate successioni illimitate di prove o di campioni verranno adottate le notazioni che seguono.

Si fa riferimento all'insieme dei tipi di campioni $\mathsf{S} := \{s_1, ..., s_n\}$ e siadotta la legge di probabilità data dalla funzione-FtR costruibile $\mathcal{P} \in \left\lceil \mathsf{S} \longmapsto \mathbb{R}_{0+} \right\rceil$ tale che, posto per ogni $i = 1, ..., n$ $p_i := \mathcal{P}(s_i)$, si ha:

$$\forall\, i = 1, ..., n \quad : \quad 0 \leq p_i \leq 1 \text{ e } \sum_{i=1}^{n} p_i = 1 \ .$$

La sequenza $\langle p_1, ..., p_n \rangle$ si chiama **distribuzione di probabilità** relativa alla legge di probabilità finita $\mathcal{P}$.

C65 b.03 Per quanto detto sugli eventi mutuamente esclusivi, per un evento $E \subseteq \mathsf{S}$ si definisce:

$$\mathcal{P}(E) \ := \ \sum_{s \in E} \mathcal{P}(s) \ .$$

C65 b.04 Sul piano della presentazione delle indagini i risultati di una singola indagine si possono anche descrivere come giunti all'indagatore da una **sorgente di segnali**.

In termini intuitivi una tale sorgente si descrive come una apparecchiatura in grado di emettere i caratteri di un suo alfabeto forniti successivamente da meccanismi poco noti all'indagatore il quale si può basare solo sulla presunzione che a ciascuno di questi caratteri è associata una propria probabilità di essere emesso.

Formalmente una sorgente di segnali è determinata da una coppia $\langle \mathsf{S}, P\rangle$ ove $\mathsf{S} = \{s_1, ..., s_n\}$ è un alfabeto finito con $n \geq 2$ e $\mathcal{P}$ è una legge di probabilità finita.
Ogni carattere o segno s_i emesso è riconosciuto come appartenente a un tipo di campione; si possono avere due generi di emissioni, quelle che possono avere emissioni con ripetizioni di tipi di caratteri e quelle sicuramente prive di ripetizioni.
Nel secondo caso le differenti probabilità di emissione dei diversi segni sono da supporre dovute alla diversità dei posizionamenti rispetto al meccanismo di scelta per ogni singola trasmissione, diversità nota al ricevente solo attraverso la funzione di probabilità.
Nel primo caso invece si possono considerare probabilità di emissione e ricezione dovute solo ai diversi numeri di campioni tra i quali viene scelto ogni singolo carattere da trasmettere. per .

C65 b.05 Diamo qualche esempio di distribuzioni di probabilità.

Si dice **distribuzione di probabilità uniforme** relativa ad S con $|\mathsf{S}| = n$ la distribuzione che assume il valore $\dfrac{1}{n}$ per ciascuno dei tipi di campioni $s_i \in \mathsf{S}$.
Evidentemente per ogni evento $E \subseteq \mathsf{S}$ si ha $\mathcal{P}(E) = \dfrac{|E|}{|\mathsf{S}|}$.
Il processo di lancio di una moneta volto a determinare se essa cade mostrando la faccia con la "testa" o quella con la "croce" riguarda uno spazio di due campioni, `testa` e `croce`; se il peso della moneta è uniformememnte distribuito (cosa che non accade per una moneta truccata) , è lecito trattare una funzione di probabilità uniforme, ponendo $\mathcal{P}(\texttt{testa}) = \mathcal{P}(\texttt{croce}) = \dfrac{1}{2}$.
Il processo di lancio di un classico dado con la forma di un cubo smussato e con densità omogenea corrisponde a un processo con spazio di 6 tipi di campioni e probabilità uniforme; per esso, con prevedoibile significato, si ha $\mathcal{P}(1) = \mathcal{P}(2) = \mathcal{P}(3) = \mathcal{P}(4) = \mathcal{P}(5) = \mathcal{P}(6) = \dfrac{1}{6}$.
Un giro di una roulette costruita correttamente, il pescaggio di una carta da un normale mazzo di 52 carte da gioco che sia stato correttamente mescolato a caso, l'estrazione del primo numero di una tombola corrispondono a processi uniformi relativi, risp., a un numero di tipi di campioni $n = 37$, $n = 52$ ed $n = 90$.

C65 b.06 I processi di indagine precedenti riguardano prove consistenti in azioni descrivibili con schemi estremamente semplice. Si hanno processi poco più elaborati quando una prova consiste nella esecuzione in un certo numero di azioni prove consistenti nella ripetizione di azioni elementari come le precedenti in condizioni indistinguibili, mantenute visibilmente inalterate.

Per studiare una serie di d successivi lanci di una moneta si deve considerare come spazio dei campioni l'insieme delle sequenze di d possibili risultati elementari `testa` e `croce`; questo equivale all'insieme delle sequenze binarie di lunghezza d e quindi ha cardinale 2^d.

Un possibile esito di una serie di d lanci di un dado è rappresentato da una d-upla di interi da 1 a 6 e una tale serie di lanci riguarda un processo con uno spazio di campioni di 6^d elementi.

Si osserva che per le precedenti sequenze di azioni è lecito pensare che l'esito di ciascuna azione non sia influenzata da quello delle precedenti.

In generale due eventi E_1 ed E_2 si dicono **eventi indipendenti** sse

$$\mathcal{P}(E_1 \cap E_2) \;=\; \mathcal{P}(E_1) \cdot \mathcal{P}(E_2) \ .$$

C65 b.07 Tornando alla sequenza di lanci di un dado, consideriamo l'evento E' corrispondente a due lanci con risultato 1 al primo lancio e l'evento E'' relativo a due lanci con risultato 1 al secondo lancio.

La probabilità di tali eventi è $\mathcal{P}(E') = \mathcal{P}(E'') = \frac{6}{36} = \frac{1}{6}$ e la probabilità di $E' \cap E''$, probabilità che si verifichino entrambi gli eventi, cioè che si abbiano due lanci con esito 1, è $P(E' \cap E'') = \frac{1}{36}$.
In effetti si tratta di due eventi indipendenti.

Consideriamo tre lanci di un dado e gli eventi E_1 riguardante i primi due lanci con esito 1 ed E_2 relativo all'ottenere 1 nel secondo e nel terzo lancio.
Lo spazio dei campioni contiene $n = 6^3 = 216$ elementi e $\mathcal{P}(E_1) = \mathcal{P}(E_2) = \frac{6}{216} = \frac{1}{36}$; $E_1 \cap E_2$ è la terna di lanci con esito 1 e quindi $\mathcal{P}(E_1 \cap E_2) = \frac{1}{216}$, valore diverso da $\mathcal{P}(E_1) \cdot \mathcal{P}(E_2) = \frac{1}{36} \cdot \frac{1}{36} = \frac{1}{1296}$.
E_1 ed E_2 quindi costituiscono un esempio di eventi non indipendenti.

C65 b.08 La tipica coppia di eventi indipendenti riguarda uno spazio dei campioni della forma $\mathsf{S} = X \times Y$, una distribuzione di probabilità uniforme, un primo evento della forma $E_1 = X' \times Y$ con $X' \subset X$ e un secondo evento della forma $E_2 = X \times Y'$ con $Y' \subset Y$.
Chiaramente $\mathcal{P}(E_1) = \frac{|X'|}{|X|}$, $\mathcal{P}(E_2) = \frac{|Y'|}{|Y|}$, $E_1 \cap E_2 = X' \times Y'$, $\mathcal{P}(E_1 \cap E_2) = \frac{|X'| \cdot |Y'|}{|X| \cdot |Y|} = \mathcal{P}(E_1) \cdot \mathcal{P}(E_2)$.

C65 b.09 Generalizziamo la nozione di indipendenza: d eventi $E_1, ..., E_d$ si dicono **eventi mutuamente indipendenti** sse per ogni scelta $E_{j_1}, ..., E_{j_h}$ tra tali eventi ($h = 2, ..., d$ e $\{j_1, ..., j_h\} \subseteq \{1, ..., d\}$), si ha

$$\mathcal{P}(E_{j_1} \cap ... \cap E_{j_h}) = \mathcal{P}(E_{j_1}) \cdot ... \cdot \mathcal{P}(E_{j_h}) .$$

Il tipico insieme di k eventi mutuamente indipendenti è dato dalla generalizzazione d-dimensionale della precedente situazione "bidimensionale".
Lo spazio dei campioni ha la forma $X_1 \times ... \times X_d$, gli eventi sono $E_1 = X'_1 \times X_2 ... \times X_d$ con $X'_1 \subset X_1$, ..., $E_d = X_1 \times ... \times X'_d$ con $X'_d \subset X_d$ e la distribuzione di probabilità è uniforme.
Chiaramente per ogni $i = 1, ..., d$ abbiamo

$$\mathcal{P}(E_i) = \frac{|X'_i|}{|X_i|} \quad \text{e} \quad \mathcal{P}(E_{j_1} \cap ... \cap E_{j_h}) = \frac{|X'_{j_1}| \cdot ... \cdot |X'_{j_h}|}{|X_{j_1}| \cdot ... \cdot |X_{j_h}|} = \mathcal{P}(E_{j_1}) \cdot ... \cdot \mathcal{P}(E_{j_h}) .$$

C65 b.10 Caso particolare importante è quello in cui tutti gli X_i coincidono: esso riguarda la replica, d volte, di una prova elementare eseguita in condizioni che l'indagatore percepisce come immutate.
Questo è il caso dei lanci successivi di una moneta o di un dado, il caso del lancio di d monete indistinguibili o di d dadi indistinguibili, il caso delle estrazioni da un'urna con reimmissione di quanto estratto o dal pescaggio da un mazzo di carte con successivo accurato rimescolamento.

Altro paradigma degli eventi indipendenti è costituito dalla emissioni di segnali da parte di una sorgente considerata priva di memoria e che opera in condizioni costanti nel tempo: per essa la probabilità di emissione dei diversi possibili tipi di segnali non cambia nel tempo e quindi tra diversi successivi segnali non si hanno correlazioni.

Per altre serie di azioni elementari accade invece che quelle che si svolgono in momenti diversi si sviluppano in circostanze diverse.
Occorre distinguere se le azioni che si sviluppano diversamente nelle fasi successive non vengano influenzate dalle precedenti oppure lo siano.
Il primo caso corrisponde a uno spazio dei campioni delle forma $X_1 \times ... \times X_d$, ciascuno dei fattori cartesiani X_i corrispondente alle condizioni che si riscontrano all'istante i e che si mantengono invariate; anche in questo caso si tratta di eventi indipendenti.

C65 b.11 Fondamentalmente diverso è il caso in cui una azione risulta influenzata dall'esito delle precedenti.
Questo si riscontra nella distribuzione di una mano di carte, nella estrazione dei 5 numeri di una determinata ruota del lotto o nella estrazione dei numeri della tombola, estrazioni notoriamente senza reimbussolamento ovvero senza reimmissione dei numeri estratti.

Esempio In una mano di poker nella quale si usano 28 carte (8, 9, 10, figure ed assi), qual'è la probabilità di avere un poker d'assi alla prima distribuzione di carte?

Lo spazio dei campioni è dato dalle possibili distribuzioni di 5 carte, cioè dai sottoinsiemi di 5 elementi dell'insieme delle 28 del mazzo; esso quindi consiste di $\binom{28}{5} = \frac{28 \cdot 27 \cdot 26 \cdot 25 \cdot 24}{5!} = 98\,280$ elementi.
Ogni evento favorevole corrisponde a un insieme di carte contenente i 4 assi e una delle 24 carte rimanenti; quindi gli eventi favorevoli sono 24.
La probabilità richiesta è quindi $\frac{24}{98\,280} = \frac{1}{4\,095} = 0.000\,244\,2...$.

C65 b.12 **Esempio** Consideriamo una sorgente che emette cifre decimali a caso, cioè con la probabilità di avere ciascuna cifra pari a $\frac{1}{10}$. Qual'è la probabilità che la quarta cifra coincida con una delle 3 precedenti?

Lo spazio dei tipi di campioni S è codificato dalle 10^4 quaterne di cifre decimali dalla 0000 alla 9999. Conviene fare rferimento al complementare $\mathsf{S} \setminus E$ dell'evento descritto e valutare la probabilità $\mathcal{P}(\mathsf{S} \setminus E)$ che la quarta cifra sia diversa da ciascuna delle prime 3. Infatti si trova facilmente che $|\mathsf{S} \setminus E| = 9^3 \cdot 10$; si trova quindi

$$\mathcal{P}(E) \;=\; 1 - \frac{9^3 \cdot 10}{10^4} \;=\; 1 - \left(\frac{9}{10}\right)^3 \;=\; 1 - 0.729 \;=\; 0.271 \ .$$

C65 b.13 I processi di indagine discreta sono utilmente trattati attraverso arborescenze [D30].

A questo proposito è opportuno riprendere le nozioni di prefisso e di arborescenza dei prefissi per le stringhe di un alfabeto $\mathsf{A} = \{a_1, ..., a_r\}$.

Date due stringhe $w, z \in \mathsf{A}^*$, si dice che w è prefisso (in senso lato) di z sse $z \in w\mathsf{A}^*$, ovvero sse $z = wx$ per qualche $x \in \mathsf{A}^*$; in questo caso si scrive $w \preceq_{\mathtt{p}} z$.
Scriviamo poi $w \prec_{\mathtt{p}} \mathbf{I}y$ per segnalare che w è prefisso proprio massimale della $y \in \mathsf{A}^*$, cioè sse $y = w\,a_j$ per qualche $a_j \in \mathsf{A}$, . ovvero sse non esiste alcuna altra stringa $v \in \mathsf{A}^*$ tale che $w \prec_{\mathtt{p}} v \prec_{\mathtt{p}} y$.
Ricordiamo anche la **arborescenza dei prefissi del monoide libero** A^* definita come

$$\Psi_{\mathtt{p}}(\mathsf{A}) \;:=\; \Big\langle \mathsf{A}^*, \{w \in \mathsf{A}^*, a_i \in \mathsf{A} :| \ \langle w, w\,a_i \rangle\}, \Big\rangle$$

avente come nodi le stringhe su A e come archi le coppie di stringhe poste in relazione dalla $\prec_{\mathtt{p}}$ **I**.
In parole povere gli archi uscenti da un nodo-stringa raggiungono tutti e soli i nodi-stringa delle quali è prefisso immediato .

C65 b.14 Come si è detto uno svolgimento di indagine si può rappresentare con una stringa $w = a_{j_1} a_{j_2} ... a_{j_d}$ i cui caratteri successivi rappresentano gli esiti delle successive prove elementari.
Questa stringa si può leggere su un cammino che parte dalla radice dell'arborescenza $\Psi_{\mathtt{p}}(\mathsf{A})$ dove A è l'alfabeto rappresentante i possibili risultati delle prove elementari effettuabili nel corso di una indagine.

Dovendo effettuare indagini esaurienti su un dato spazio dei campioni, occorre precisare una **strategia di indagine**, cioè determinare una procedura che consenta di scegliere quali prove elementari effettuare per ottenere nuove informazioni utili.

Una strategia di indagine corrisponde alla sottoarborescenza di $\Psi_{\mathsf{p}}(\mathsf{A})$ i cui cammini massimali costituiscono i previsti possibili svolgimenti di indagine.

C65 b.15 Ricordiamo che, stabilito un ordine totale $\leq$ per i caratteri dell'alfabeto A, $\Psi_{\mathsf{p}}(\mathsf{A})$ diventa un'arborescenza distesa se, si ordinano i figli di ogni nodo w secondo l'ordine assegnato agli elementi di A, wa_1, wa_2, ..., wa_r, cioè secondo l'ordine lessicografico $\leq_{lxg}$ (il quale in questo ambito coincide con l'ordine secondo lunghezza-lessicografico $\leq_{llx}$).

Quando si può prescindere dal significato dei singoli caratteri di A ma interessa solo che siano condivisibilmente distinguibili, si può assumere che essi siano gli interi 0, 1, ..., $r-1$ e ritenerli implicitamente ordinati secondo valore.

C65 b.16 Vediamo alcune semplici arborescenze di indagine corrispondenti ai processi visti in precedenza.

La più semplice di esse è l'arborescenza relativa a d lanci di una moneta. La serie di lanci si può considerare finalizzata a individuare una delle possibili d-uple di esiti. Le successive prove non sono influenzate dalle precedenti e ogni lancio ha due possibili esiti: quindi abbiamo un'arborescenza in cui ogni padre ha due figli (`testa` e `croce`). Ogni cammino massimale deve avere lunghezza d.

I processi che si svolgono in condizioni che mutano nel tempo sono rappresentati da arborescenze nonuniformi.

Il caso della estrazione dei 5 numeri del lotto è rappresentabile da un'arborescenza con una prima arborescenza-1 con 90 archi, con 90 arborescenze-1 di 89 archi che finiscono al livello 2 e rappresentano il secondo numero estratto, con $90 \cdot 89 = 8\,010$ arborescenze-1 di 88 archi che finiscono al livello 3, con $90 \cdot 89 \cdot 88 = 704\,880$ arborescenze di 87 archi verso il livello 4 e con $90 \cdot 89 \cdot 88 \cdot 87 = 61\,324\,560$ arborescenze-1 di 86 archi verso i $90 \cdot 89 \cdot 88 \cdot 87 \cdot 86 = 5\,273\,912\,160$ nodi foglia.

Lo spazio dei campioni è in biiezione con l'insieme dei cammini massimali di questa arborescenza ovvero con l'insieme delle sue foglie, evidentemente individuabili con le disposizioni senza ripetizioni dei primi 90 interi positivi.

C65 c. nozione di codice

C65 c.01 Al termine "codice" si attribuiscono molti significati, anche nell'ambito della sola matematica. Qui lo utilizzeremo come sinonimo di linguaggio, ma limitatamente allo studio dei linguaggi rivolto alle attività di trasmissione e di valutazione delle informazioni.

C65 c.02 Consideriamo un intero r e un alfabeto A di r caratteri; chiamiamo **precodice di arietà** r o precodice r-ario ogni linguaggio su A e chiamiamo **parole di codice** le sue stringhe.

Risulta opportuno mettere l'alfabeto Adi r caratteri in corrispondenza biunivoca con $\mathbb{Z}_r$ in quanto anello finito per la possibilità di utilizzare tecniche algebriche che lo riguardano.
Queste sono particolarmente efficaci quando $\mathbb{Z}_r$ è un campo, ossia quando r è un numero primo o una potenza di un numero primo.

Particolarmente importanti sono i **precodici binari**, codici sull'alfabeto $\{0,1\}$, in quanto essi possono essere oggetto di rielaborazioni e di trasmissioni eseguibili con i dispositivi hardware e software digitali attualmente disponibili, strumenti caratterizzati da elevata efficienza e versatilità e in grado di trattare ogni forma di informazione razionalizzabile mediante sequenze di bits.

C65 c.03 Una operazione molto richiesta per la elaborazione delle informazioni consiste nel sostituire le stringhe su un alfabeto $\mathsf{S} = \{s_1, ...\}$, finito o infinito, con stringhe su un secondo alfabeto $\mathsf{A} = \{a_1, ..., a_r\}$ in modo da essere di trasformare ciascuna di queste stringhe su A nella stringa su SSs sua corrispondente.
Una tale trasformazione risulta utile soprattutto quando il cardinale di S è sensibilmente maggiore di r (o quando si ha un S è illimitato); in particolare risulta utile quando $r = 2$.

Formalmente ci serviamo di una biiezione $\kappa \in \lceil \mathsf{S} \lhd\!\!\longrightarrow\!\!\rhd \mathsf{C} \rceil$ con $\mathsf{C} \subset \mathsf{A}^+$ chiamata **funzione di codifica** e la sua estensione per giustapposizione, che, potendo semplificare, denoteremo ancora con κ.

Ogni stringa $s_{\pi_1}...s_{\pi_h}$ su S può essere codificata nella $\kappa(s_{\pi_1})...\kappa(s_{\pi_h}) \in \mathsf{A}^+$.
Si osserva che una tale funzione κ è un morfismo di monoidi.
Chiameremo allora **alfabeto sorgente** l'insieme S, precodice il linguaggio C e **schema di codifica** la terna $\langle \mathsf{S}, \mathsf{C}, \kappa \rangle$.

C65 c.04 Tra gli schemi di codifica ricordiamo lo schema ASCII, secondo il quale un insieme di un paio di centinaia di simboli (cifre decimali, lettere maiuscole e minuscole, simboli matematici elementari, simboli utilizzati per il controllo della stampa e delle trasmissioni, ...) vengono espressi mediante ottetti binari.

Conviene ricordare anche il codice **Unicode** costituito da gran parte delle sequenze di 16 bits e da varie altre sequenze binarie di maggiore lunghezzaviene proposto da varie autorità internazionali come veicolo per ogni genere di comunicazione leggibile e ormai viene ampiamente adottato come strumento in grado di esprimere tutti i simboli utilizzati da centinaia di lingue naturali e da una grande quantità di convenzioni per prodotti di largo uso nelle attività industriali, commerciali e pubblicistiche.

C65 c.05 Per le sostituzioni, ovvero per gli schemi di codifica, è molto importante la proprietà di **decifrabilità univoca**, cioè la possibilità che da ogni sequenza di parole di codice si possa ottenere una sola stringa sull'alfabeto sorgente S.

Più formalmente un precodice C si dice **precodice unicamente decifrabile**, ovvero si dice che C è un **codice**, sse per ogni stringa $a_{\pi_1}...a_{\pi_h} \in \mathsf{A}^*$ esiste al più una sequenza di parole codice $c_{\alpha_1}...c_{\alpha_d}$ tale che

$$c_{\alpha_1}...c_{\alpha_d} = a_{\pi_1}...a_{\pi_h}.$$

Ad esempio il precodice binario $\{0, 01, 001\}$ non è unicamente decifrabile, in quanto 0001 si può scandire come $0|001$ oppure come $0|0|01$ (qui utilizziamo il segno " | " per marcare la separazione tra due parole codice consecutive).

Si osserva anche che una parola del codice precedente si può ottenere come giustapposizione di altre due: questa proprietà chiaramente è in conflitto con la decifrabilità unica.

C65 c.06 Viceversa è unicamente decifrabile il precodice $\{0, 10, 110\}$: consideriamo infatti il seguente accettatore-trasduttore

stato		0		1	
s_0	$\mp$	$0/$	s_0	$\sqcup$	s_1
s_1		$10/$	s_0	$\sqcup$	s_2
s_2		$110/$	s_0	**err**	

Una qualsiasi stringa binaria che gli viene sottoposta o viene scandita come sequenza di parole del precodice dato, oppure viene rifiutata in quanto presenta tre o più 1 consecutivi, oppure in quanto si conclude con 1 o con 11.

C65 c.07 Un precodice costituito da parole della stessa lunghezza è evidentemente un codice e viene chiamato **codice a lunghezza fissa**; in caso contrario viene detto **precodice a lunghezza variabile**.

I codici ASCII-7 e ASCII-8, come pure molti altri codici utilizzati in telegrafia e con i computers delprimo periodo (fino al 1970 circa) e ora piuttosto desueti (come Morse, Baudot, codice biquinario, BCD ed EBCDIC) sono codici a lunghezza fissa.

Il codice Unicode è costituito in gran parte da parole codice di 16 bits, ma anche da parole che si riconoscono essere costituite da 3 o più ottetti di bits; inoltre nei messaggi Unicode si possono riconoscere stringhe costituite da ottetti compatibili con ASCII-8 e altre codifiche a lunghezza fissa; questa proprietà è stata voluta per consentire di utilizzare codifiche precedenti con strumenti in grado di trattare tutti i testi Unicode.

Evidentemente ogni precodice a lunghezza fissa t è unicamente decifrabile: per la decifratura di una parola w basta individuare i suoi successivi fattori di lunghezza t (se una tale fattorizzazione non è attuabile w viene riconosciuta come parola non codificata).

Per contro un tale codice può essere poco efficiente innanzi tutto in termini di lunghezza delle sequenze che codificano gran parte dei messaggi.
Va anche osservato che il risparmio di spazio per la memorizzazione di testi codificati comporta risparmio di tempo per la loro trasmissione. stante la odierna importanza della memorizzazione e della trasmissione digitale delle informazioni, gli accennati risparmi costituiscono una questione di elevata importanza pratica.

Per questi risparmi si può subito osservqare che un lungo testo su S potrebbe essere codificato con un numero inferiore di caratteri dell'alfabeto A' mediante uno schema di codifica $\langle \mathsf{S}, \mathsf{C}', \kappa' \rangle$ se κ' si preoccupa di associare le parole di codice più corte ai caratteri sorgente più frequenti.

C65 c.08 Un codice si dice **codice immediatamente decifrabile** sse consente di decifrare ogni sua parola di codice in un testo trasmesso non appena è stata ricevuto l'ultimo dei suoi caratteri.

Un linguaggio si dice **linguaggio a prefissi** sse nessuna delle sue stringhe è prefisso di un'altra; se un tale linguaggio viene usato come codice viene chiamato **codice a prefissi**.

I codici a prefisso si trattano naturalmente facendo riferimento ad arborescenze distese. Infatti un codice è un linguaggio le cui stringhe sono non confrontabili rispetto alla relazione d'ordine $\preceq_{\mathtt{p}}$, "essere prefisso di o coincidere con".

C65 c.09 Conviene vedere le parole di un codice a prefissi sull'arborescenza infinita dei prefissi $\Psi_{\mathtt{p}}(\mathsf{A})$.

Prop. Un linguaggio sull'alfabeto A è codice a prefissi su A sse le sue stringhe costituiscono una anticatena di $\Psi_{\mathtt{p}}(\mathsf{A})$, cioè sse non si trovano mai due di esse su uno stesso cammino ∎

Si dice **arborescenza di un codice a prefissi** la sottoarborescenza della $\Psi_{\mathtt{p}}(\mathsf{A})$ ottenuta considerando solo i suoi cammini dalla radice $\sqcup$ ai nodi delle parole codice. Questa arborescenza fornisce una efficace raffigurazione del codice a prefissi. Esempi:

C65 c.10 Prop. Un linguaggio a prefissi C è un codice (unicamente decifrabile).

Dim.: È possibile costruire un accettatore-trasduttore che effettua la decifrazione a partire dall'arborescenza delle parole di codice, cioè della sottoarborescenza di $\Psi_{\mathtt{p}}(\mathsf{A})$ ottenuta limitandosi alle parole di C. Per questo basta modificare questa arborescenza rimpiazzando ogni arco che porta a una sua foglia $\langle x, xa\rangle$ con un arco bietichettato $\langle x, a, xa, \sqcup\rangle$ ∎
I codici a prefisso coincidono con i codici immediatamente decifrabili.

C65 c.11 Riassumendo: i codici a prefissi possono essere finiti o infiniti.
I codici a prefissi finiti sono in biiezione con le anticatene di $\Psi_{\mathtt{p}}(\mathsf{A})$; tutte queste anticatene posseggono estensioni massimali.
Alle anticatene massimali di $\Psi_{\mathtt{p}}(\mathsf{A})$ sono associati biunivocamente i codici massimali, codici che possono essere particolarmente vantaggiosi.
Tutti gli altri codici a prefissi finiti sono in biiezione con i sottoinsiemi propri non vuoti delle anticatene massimali.

C65 c.12 Dato un codice, ogni suo sottoinsieme nonvuoto è un codice; infatti è impossibile trovare una bifattorizzazione di qualche parola su A^* servendosi solo di una parte delle parole di un codice.

Interessano quindi prevalentemente i codici massimali, codici privi di sovrainsiemi che siano codici essi stessi.

Non è detto che un codice possegga un codice massimale finito; in particolare i codici binari a virgola sono sottoinsiemi di codici infiniti della forma 0^*1.

C65 c.13 Prop. Un codice a prefissi f-ario finito è massimale sse la sua arborescenza risulta uniforme-f.

∎

Quindi ogni codice a prefissi finito f-ario si può ricavare dalle rotte per un determinato sottoinsieme delle foglie di un'arborescenza uniforme-f.

C65 c.14 Alle foglie di un'arborescenza uniforme-r, ovvero alle parole di un codice a prefissi massimale, si può associare una **distribuzione di probabilità canonica** definita assegnando ad ogni foglia di altezza h la probabilità $\dfrac{1}{r^h}$.

Dimostriamo che $\displaystyle\sum_{i\in foglie} \frac{1}{r^{h_i}} \;=\; 1.$

Per la somma sui figli di un nodo ν di altezza $h-1$ si ha

$$\sum_{figli\ di\ \nu} \frac{1}{r^{t_h}} \;=\; \frac{1}{r^{h-1}}\ .$$

Quindi la suddetta somma si riduce alla somma sugli r figli della radice:

$$\sum_{i=1}^{r} \frac{1}{r} \;=\; 1\ .$$

C65 c.15 Si osserva che un codice unicamente decifrabile non può presentare un numero eccessivo di parole corte. Più precisamente vale il seguente fatto.

Teorema (**teorema di McMillan 1957**)
Consideriamo il codice r-ario $\mathsf{C} = \{c_1, ..., c_q\}$ e le corrispondenti lunghezze $t_1 := |c_1|, ..., t_q := |c_q|$.
Se C è unicamente decifrabile, allora deve essere

$$\sum_{k=1}^{q} \frac{1}{r^{t_k}} \leq 1.$$

C65 c.16 Si dice **codice istantaneo** un codice che per qualsiasi messaggio ricevuto consente di interpretare ogni parola codice immediatamente dopo la sua ricezione.
Evidentemente ogni codice istantaneo è un codice unicamente decifrabile.

Un codice è istantaneo sse possiede la proprietà del prefisso.

C65 c.17 Teorema (**teorema di Kraft 1958**)
Esiste un codice di radice r $C = \{c_1, c_2, ..., c_q\}$ con le lunghezze delle parol codice ℓ_1, ℓ_2, ... , ℓ_q sse è soddisfatta la proprietà di Kraft

$$\sum_{k=1}^{q} \frac{1}{r^{\ell_k}} \leq 1.$$

Si dice **codice massimale** un codice che non è contenuto in alcun altro codice.

Si dice codice istantaneo massimale un codice che non è contenuto in alcun altro codice istantaneo.

Un codice istantaneo è massimale sse per esso nella disuguaglianza di Kraft vale il segno uguale.

Consideriamo un codice istantaneo avente m come massima lunghezza delle parole codice; se esso non è massimale può essere ampliato rimanendo istantaneo con una parola codice di lunghezza m.

Ad ogni codice unicamente decfrabile si può associare un codice istantaneo caratterizzato dalla stessa sequenza di lunghezze di parole codice.

Testo fruibile in https://www.mi.imati.cnr.it/alberto/ e https://arm.mi.imati.cnr.it/Matexp/matexp_main.php