TECNICA GREEDY: MASSIMO NUMERO DI INTERVALLI DISGIUNTI

Questa parte non è presente in nessuno dei testi adottati.

Il problema

Dato un **insieme di intervalli**, in generale **non** tutti **disgiunti** fra di loro, trovarne un **sottoinsieme costituito da intervalli tutti disgiunti e tale che il numero di intervalli sia il massimo**.

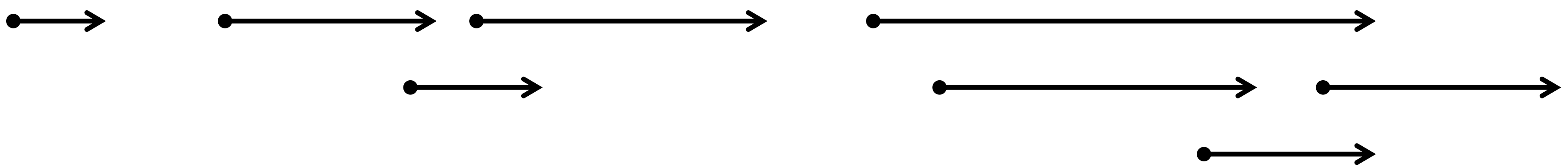
Esempio: Il problema *Nimbus* dalle Olimpiadi di Informatica.

Al celebre maghetto Harry Potter è stata regalata una scopa volante modello Nimbus3000 e tutti i suoi compagni del Grifondoro gli chiedono di poterla provare. Il buon Harry ha promesso che nei giorni a venire soddisferà le richieste di tutti, ma ogni ragazzo è impaziente e vuole provare la scopa il giorno stesso.

Ognuno propone a Harry un intervallo di tempo della giornata durante il quale, essendo libero da lezioni di magia, può fare un giro sulla scopa, e per convincerlo gli offre una fantastica caramella*Tuttigusti+1.* Tenendo presente che una sola persona alla volta può salire sulla Nimbus3000 in ogni istante di tempo, Harry decide di soddisfare, tra tutte le richieste dei ragazzi, quelle che gli procureranno la massima quantità di caramelle (che poi spartirà coi suoi amici Ron e Hermione). Aiutalo a trovare la migliore soluzione possibile.

Importante: Harry non conosce ancora magie di scheduling.

Quale politica usare?



Scandiamo l’insieme di intervalli: ad ogni passo aggiungiamo all’insieme soluzione un nuovo intervallo, **disgiunto** da quelli già aggiunti nei passi precedenti.

Secondo quale **criterio** scegliamo l’intervallo (ovviamente fra quelli disgiunti dai precedentemente inclusi nella soluzione) da aggiungere ogni volta, per fare in modo che alla fine il numero di intervalli sia **massimo**?

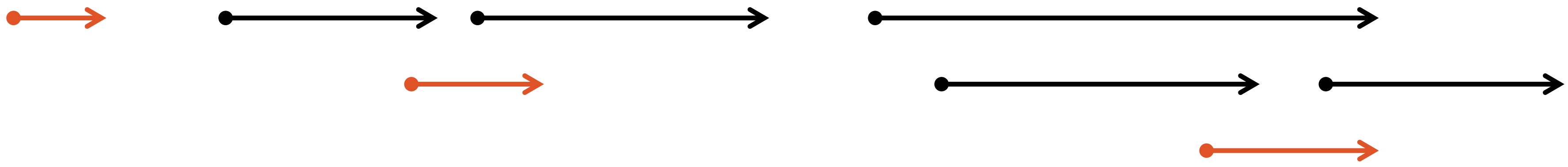
Nota: una sequenza di intervalli disgiunti si definisce massimale se contiene il **massimo numero possibile di intervalli**.

Quale politica usare?

Proviamo con criteri diversi.

- scegliamo ogni volta **l'intervallo che comincia prima**: l'insieme finale ha **4** elementi;

Quale politica usare?



Proviamo con criteri diversi.

- scegliamo ogni volta l'intervallo che comincia prima: l'insieme finale ha 4 elementi;
- scegliamo ogni volta **l'intervallo più corto**: l'insieme finale ha **3** elementi;

Quale politica usare?

Proviamo con criteri diversi.

- scegliamo ogni volta l'intervallo che comincia prima: l'insieme finale ha 4 elementi;
- scegliamo ogni volta l'intervallo più corto: l'insieme finale ha 3 elementi;
- scegliamo ogni volta l'intervallo **che finisce prima**: l'insieme finale ha **5** elementi.

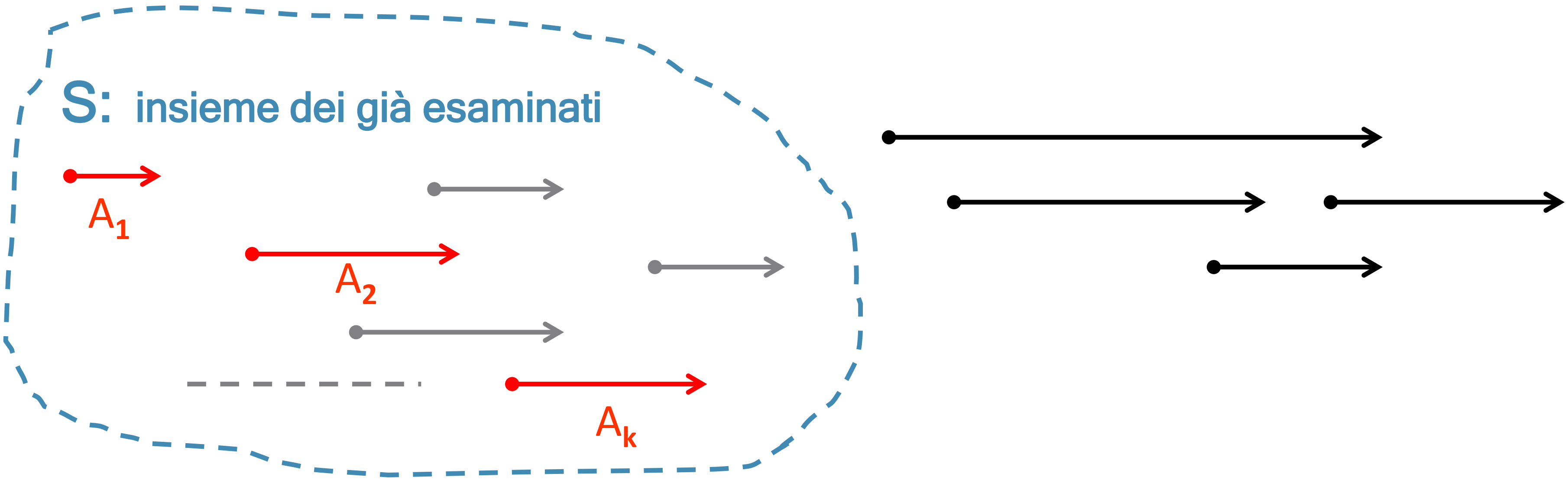
Si dimostra che l'ultimo è effettivamente il criterio corretto, che assicura **l'ottimalità** della soluzione (cioè che il numero di intervalli scelti sia massimo).

L'algoritmo

1. ordina l'insieme degli intervalli in una **sequenza S ordinata** (in maniera non decrescente) **secondo l'istante finale** (cioè, geometricamente, secondo l'estremo destro del segmento);
2. inizializza la soluzione **Sol** come **sequenza vuota**;
3. scandisci S nell'ordine, e per ogni suo elemento A:
   1. Se A **inizia dopo la fine dell'ultimo elemento di Sol**, **aggiungilo** al fondo di **Sol**;
   2. Altrimenti non aggiungerlo (implicitamente: scarta A).

COMPLESSITÀ: la complessità è quella dell'ordinamento del passo 1 (**O(n log n)**) più la scansione di S al passo 3 (**O(n)**). Quindi la complessità totale è **O(n log n)**.

Dimostrazione di correttezza



Situazione a un generico passo intermedio (Invarianti), dove **S** è **l'insieme di tutti gl'intervalli finora esaminati**:

1. la sequenza $\mathbf{A_1, A_2, \ldots, A_k}$ di intervalli disgiunti già scelti è, per l'insieme **S**, una sequenza **massimale**, e inoltre
2. fra le sequenze massimali **è quella che finisce per prima** (cioè il cui istante finale dell'ultimo intervallo è il minimo);
3. ogni intervallo $\notin$**S termina dopo** (o non prima) della fine di ogni intervallo $\in$**S**.

Nomi invarianti

Diamo agli invarianti dei nomi lievemente più mnemonici:

Max: la sequenza $A_1, A_2, \ldots, A_k$ di intervalli disgiunti già scelti è, per l'insieme S, una **sequenza massimale**;

PrimaMax: la sequenza $A_1, A_2, \ldots, A_k$ è, fra le sequenze massimali, quella **che finisce per prima** (cioè il cui istante finale dell'ultimo intervallo è il minimo);

PrimaVisti ogni intervallo $\in S$ (anche non appartenente alla sequenza massimale) **termina prima** della fine di qualunque intervallo $\notin S$;

cioè gli estremi destri degli intervalli già esaminati (inclusi o no nella sequenza massimale) sono tutti a sinistra di (cioè precedono) tutti gli estremi destri degli intervalli ancora da esaminare.

Osservazioni

- Noi vogliamo dimostrare che, all’uscita dal ciclo, la sequenza **A_1, A_2, …, A_m** di intervalli disgiunti scelti è, per l’insieme di tutti gli intervalli dati, una sequenza **massimale**.
- Quindi vogliamo dimostrare che **Max** è un invariante dell’algoritmo.
- Per dimostrare ciò dobbiamo «rafforzare l’induzione»: infatti, come si vede nella slide precedente, **Max** non è il solo invariante, ma sono invarianti anche **PrimaMax** e **PrimaVisti**.
- Tutti e tre i punti sono infatti necessari per dimostrare il passo induttivo, e quindi per dimostrare che all’uscita dal ciclo vale in particolare l’invariante **Max**, che è quello che ci interessa.

Dimostrazione di correttezza (per induzione): la base

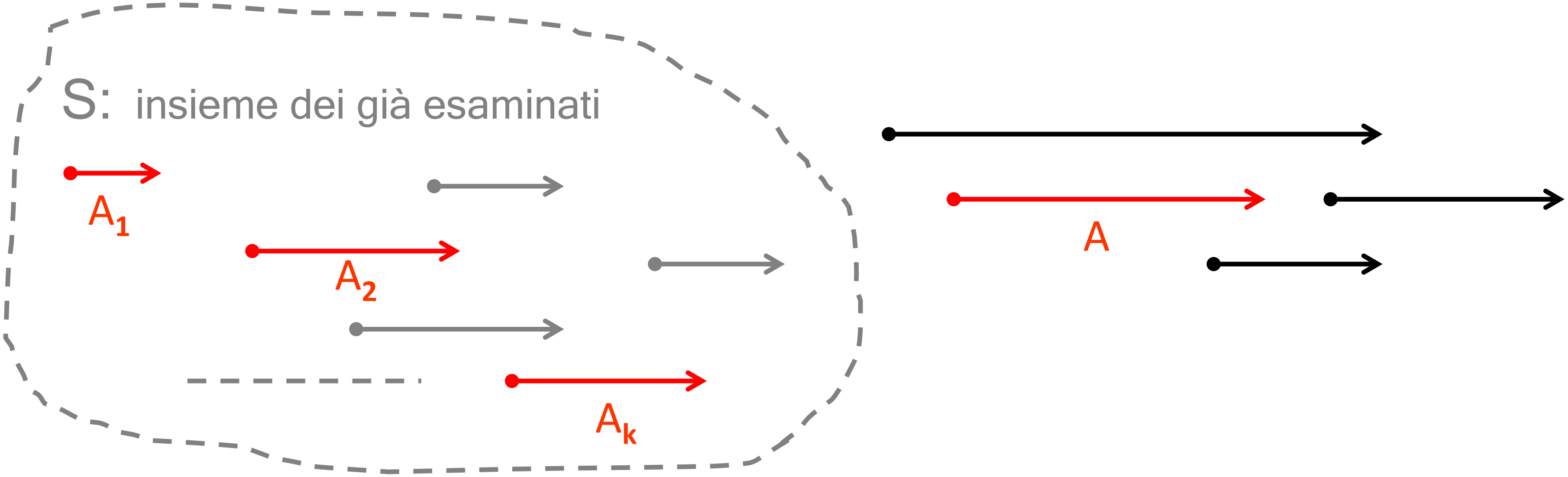
Inizialmente **S è l'insieme vuoto** e la sequenza massimale per S è la sequenza vuota: gli invarianti sono ovviamente veri.

Max: la sequenza vuota è una sequenza massimale considerato $S = \emptyset$

PrimaMax: la sequenza vuota è l'unica possibile, quindi anche quella che finisce prima

PrimaVisti: ogni intervallo finisce dopo la sequenza vuota.

Dimostrazione di correttezza: il passo.

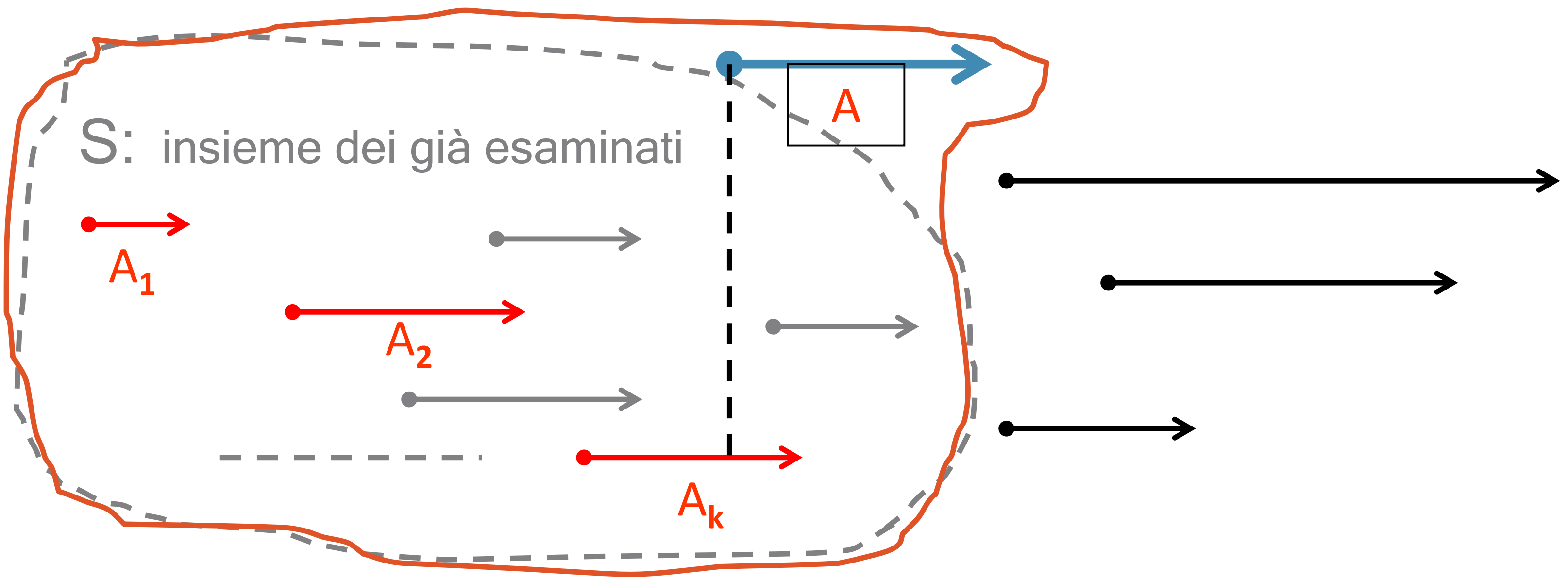


Sia **A** l'intervallo **che termina prima** fra quelli ancora da esaminare (cioè fra quelli $\notin$ **S**).

Consideriamo allora l'insieme **S' ≡ S ∪ {A}** e cerchiamo di stabilire qual è per **S'** la **sequenza massimale (Max)** che **finisce per prima (PrimaMax)**.

(Banalmente, se scelgo A in questo modo, **PrimaVisti** continuerà a valere).

Dimostrazione del passo: caso 1.

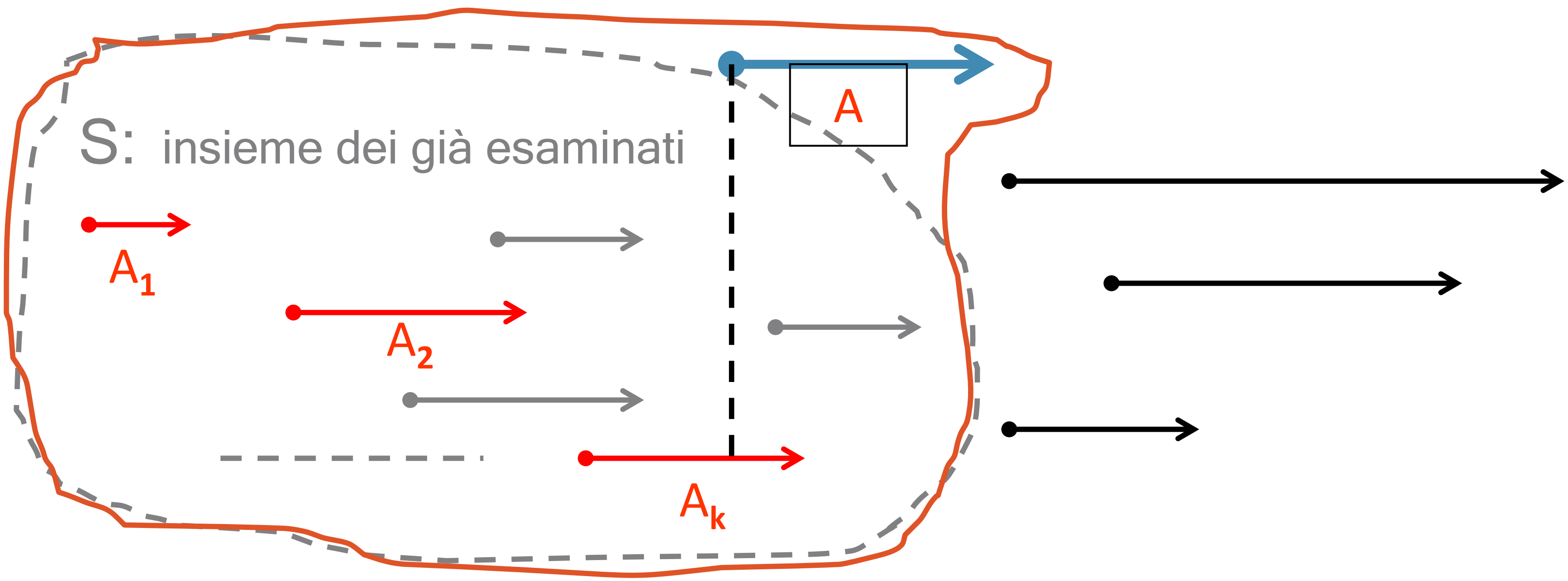


caso 1) **A inizia prima della fine della sequenza A_1, A_2, …, A_k**.

Per **ipotesi induttiva** (***PrimaMax***) ogni sequenza massimale di **S** termina dopo (o in concomitanza con) la fine di **A_k** e quindi interseca **A**.

Allora ogni sequenza avente **A** come ultimo elemento non può contenere più di **k-1** elementi di **S.**

Dimostrazione del passo: caso 1.

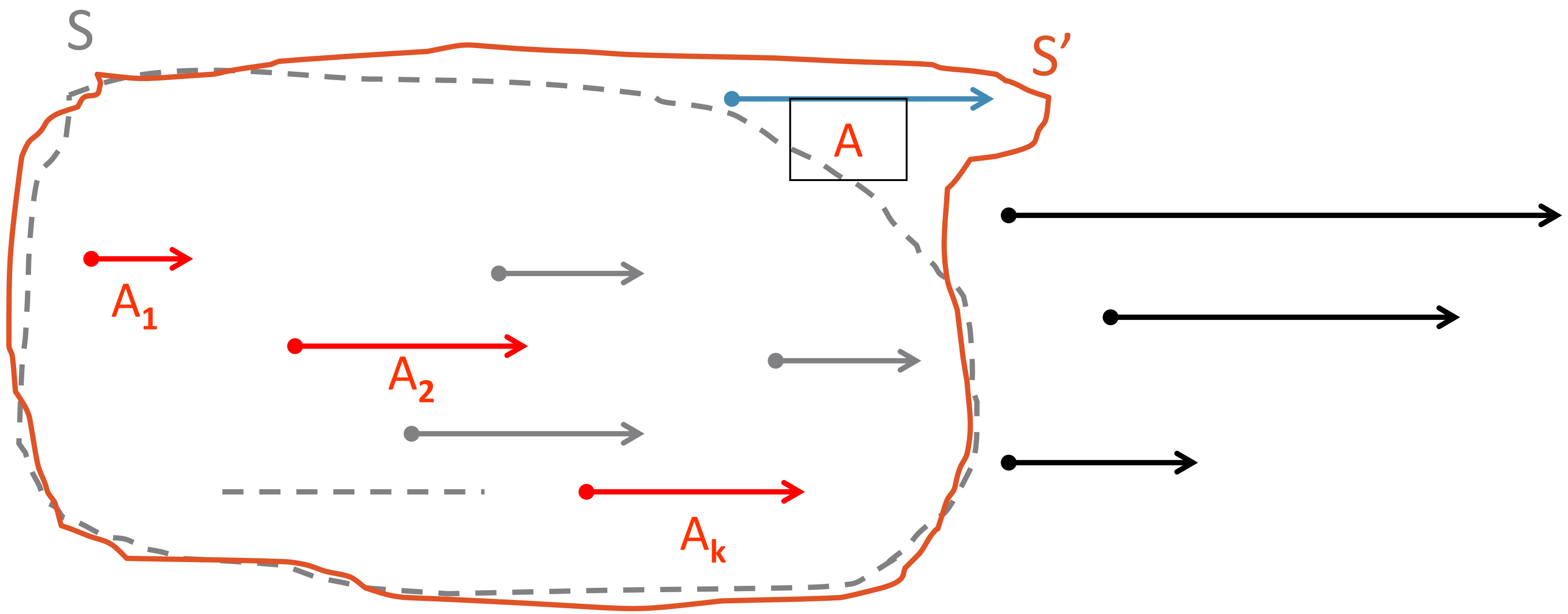


caso 1) **A inizia prima della fine della sequenza A_1, A_2, …, A_k .**

…

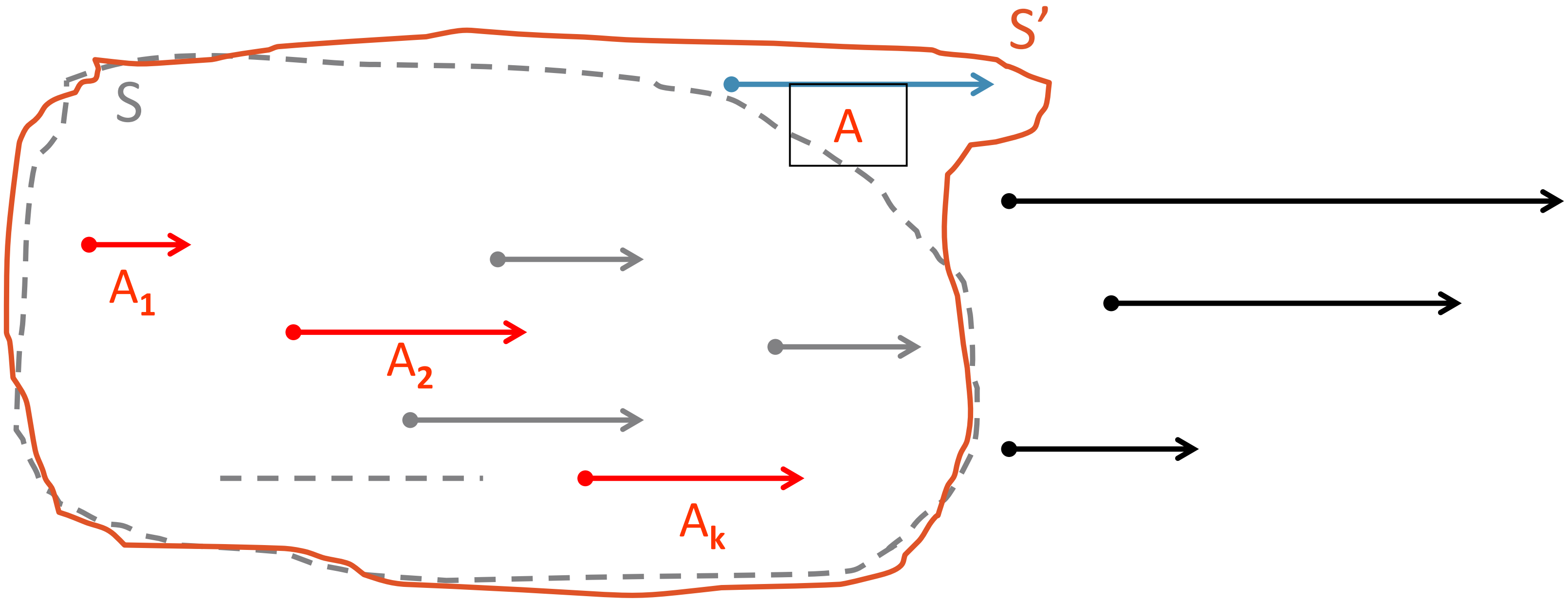
Quindi **ogni sequenza avente A come ultimo elemento non può avere più di k elementi**, e A_1, A_2, …, A_k , avendo k elementi, è una sequenza massimale anche per $S' \equiv S \cup \{A\}$; quindi, **non inserendo A nella soluzione, l'invariante *Max* si mantiene**.

Dimostrazione del passo: conclusione del caso 1.



Inoltre, per ipotesi induttiva (***PrimaVisti***), **tutti gli elementi di S**, fra i quali c'è A_k, **hanno estremi destri che precedono gli estremi destri di tutti gli elementi $\notin$S**, quindi anche di A; dunque la fine di A_k precede quella di A; pertanto A_1, A_2, …, A_{k-1}, A, se anche è una sequenza massimale per S’, termina dopo la sequenza massimale A_1, A_2, …, A_k; **pertanto per S’ la seq. massimale che finisce prima è ancora A_1, A_2, …, A_k.**

Dimostrazione del passo: conclusione del caso 1.



Dunque togliendo **A** dagli elementi ancora da esaminare ma non inserendolo nella soluzione anche l'invariante ***PrimaMax*** si mantiene.

Nella dimostrazione precedente (passo, caso 1), osserva:

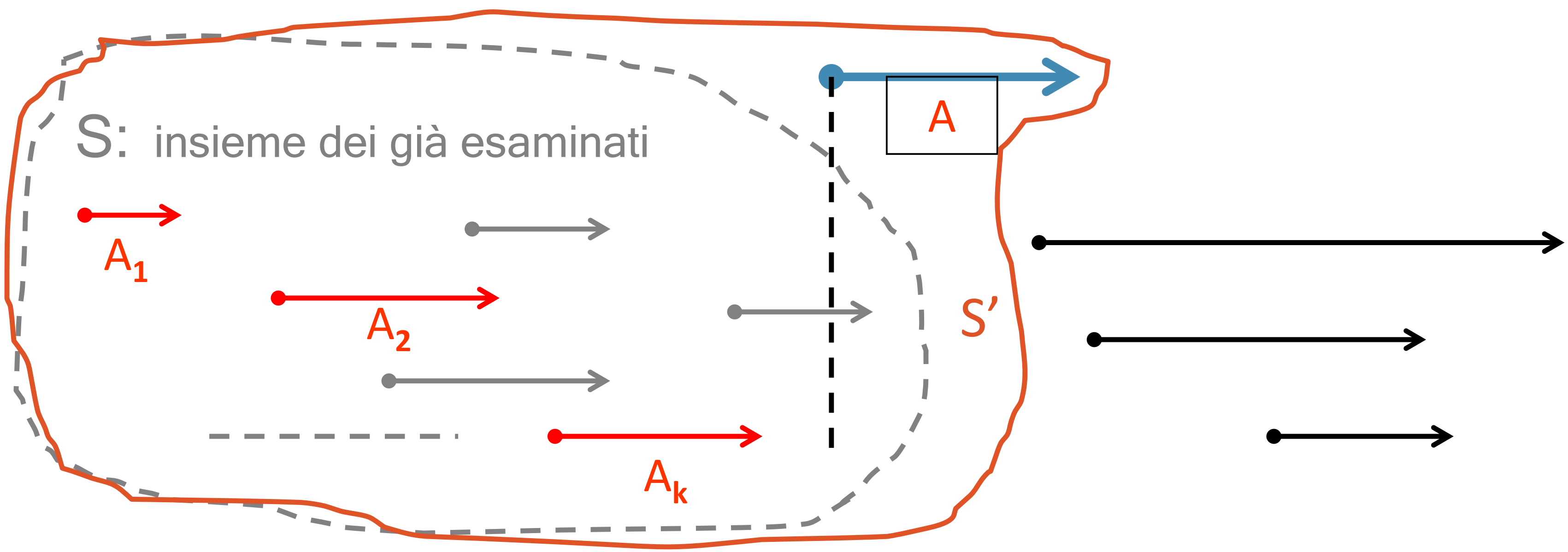
- **PrimaMax** serve, come ipotesi induttiva, per dimostrare che, dopo un'esecuzione del corpo del ciclo, **nel caso 1** vale ancora l'invariante ***Max***, cioè la sequenza $\mathbf{A_1, A_2, \ldots, A_k}$ è ancora una sequenza massimale anche per $\mathbf{S \cup \{A\}}$.
- Ma per poter dimostrare che dopo un'esecuzione del corpo del ciclo nel caso 1 vale ancora l'invariante ***PrimaMax***, è necessario assumere, come ipotesi induttiva, ***PrimaVisti***.
- Si noti infine che ***PrimaVisti*** continua a valere a causa del criterio di scelta di **A**: **A** è l'intervallo che finisce per primo fra tutti quelli $\notin \mathbf{S}$; esso potrà venire incluso oppure no nella sequenza massimale, ma in ogni caso rimane vero che tutti gli estremi destri degli elementi esaminati (cioè $\in \mathbf{S \cup \{A\}}$) precedono tutti quelli degli elementi ancora da esaminare.

Ricapitolando

Abbiamo quindi complessivamente dimostrato che:

- **se** immediatamente **prima** di una esecuzione del corpo del **ciclo** valgono tutti e tre gli **invarianti**,
- alla fine di essa, **nel caso 1**, **valgono ancora tutti e tre**.

Dimostrazione del passo: caso 2.



caso 2) **A inizia dopo la fine della sequenza A_1, A_2, …, A_k**.

Allora la sequenza **A_1, A_2, …, A_k, A**:

- è per **S'** una **sequenza massimale**, perché una seq. massimale per **S'** non può avere più di **k+1** elementi (altrimenti per **S** ci sarebbe una sequenza massimale con più di **k** elementi);
- è per **S'** la sequenza massimale **che finisce prima**, perché in **S'** non ci sono sequenze massimali non includenti **A**.

Dimostrazione del passo, caso II: *PrimaVisti*.

Nella slide precedente abbiamo dimostrato che **nel caso 2**, grazie a ***Max*** usato come ipotesi induttiva, dopo un'esecuzione del corpo del ciclo valgono ancora ***Max*** e ***PrimaMax***.

Infine, anche **nel caso 2** l'invariante ***PrimaVisti*** si conserva semplicemente a causa del criterio di scelta di **A.**

Abbiamo quindi dimostrato che **anche nel caso 2**
i tre invarianti si mantengono.

Conclusione

I tre invarianti sono quindi **validi** dopo ogni iterazione, in particolare anche **all'uscita finale dal ciclo**;

Quindi **Max** vale alla fine del ciclo;

La **sequenza finale è** perciò **una sequenza massimale** (anzi, se esistono più sequenze massimali, **è quella che finisce per prima**).

Cosa devo aver capito fino ad ora

- Il problema del massimo numero di intervalli disgiunti
- La tecnica greedy che lo risolve
- La dimostrazione di correttezza della tecnica greedy

…se non ho capito qualcosa

- Alzo la mano e chiedo
- Ripasso sul libro
- Chiedo aiuto sul forum
- Chiedo o mando una mail al docente