

# SokoBOT: a Sokoban solver attempt

Progetto di Corso di Fondamenti di Intelligenza Artificiale

<https://github.com/NANISOREH/SokoBOT>

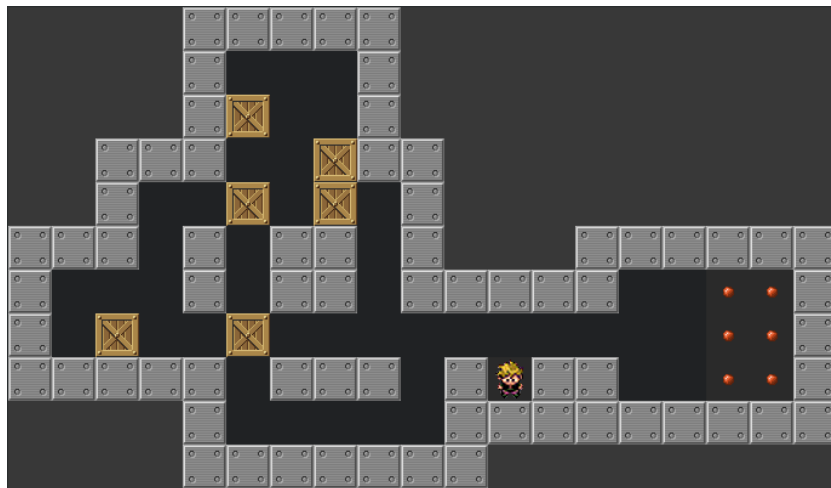
## Indice

1 – Introduzione.....	3
2 – Sokoban solver come agente intelligente.....	5
2.1 – Ambiente e Sensori.....	5
2.2 – Attuatori e misure di Prestazione.....	7
3 – Sokoban come problema di ricerca.....	9
3.1 – Stati e transizioni.....	9
3.2 – Formalizzazione del problema.....	10
3.3 – Schemi di espansione dei nodi nell'albero di ricerca.....	11
4 – Enter SokoBOT.....	13
4.1 – Elementi di difficoltà.....	13
4.2 – Rappresentazione e trasposizione degli stati.....	14
4.3 – Algoritmi di ricerca non informata.....	15
4.4 – Euristiche di stima e algoritmi di ricerca informata.....	17
4.5 – Deadlock detection.....	18

# 1 – Introduzione

Sokoban - traducibile approssimativamente con “magazziniere” - è un rompicapo creato e commercializzato in Giappone all’inizio degli anni ‘80. Il giocatore, che per semplicità - e affetto - chiamerò Sokoban da qui in avanti, si trova all’interno di un magazzino insieme ad un certo numero di casse che gli è consentito spingere, ma non tirare. Nel magazzino troverà degli specifici punti obiettivo presso cui dovrà spingere le casse. Il numero degli obiettivi è pari a quello delle casse.

L’obiettivo del gioco, come facilmente intuibile, è spingere ognuna delle casse in un punto obiettivo, preferibilmente nel minor numero possibile di mosse. Un livello tipico del gioco è costruito come un labirinto in cui i muri sono disposti in modo da costringere Sokoban a valutare attentamente l’ordine e la direzione delle proprie “spinte” alle casse. Considerando che Sokoban non può attraversare né scalare i muri, in ogni momento può fare una di queste due cose: muoversi in una posizione vuota del livello, oppure, se si trova in prossimità di una cassa, spingerla. Non gli è possibile spingere più casse “in fila indiana”: una spinta è possibile solo se, oltre la cassa che si vuole spingere, c’è uno spazio libero.



*Livello #1 del set originale di Sokoban*

Il set di regole del gioco è incredibilmente semplice. Probabilmente, se avessi potuto integrare un browser game in questo documento, non avrei nemmeno avuto bisogno di illustrarlo. Eppure, il problema computazionale ad esso sotteso è eccezionalmente complesso, e questo, unito alla facilità con cui si presta a fare da toy problem per modellare problemi reali di robotica e logistica, lo ha reso un grande classico nel campo della ricerca sull’IA. Dimostrabilmente, il problema di risolvere Sokoban appartiene alla classe di complessità NP-hard. Avrò modo di entrare nel dettaglio degli specifici elementi di difficoltà nella ricerca di soluzioni per un livello.

Questo lavoro di analisi del problema di ricerca sarà corredato da una suite di algoritmi risolutivi implementati in Java, integrati in un’applicazione grafica JavaFX che permetta di configurarli,

lanciarli e osservarne i risultati su un campione di livelli del gioco. L'obiettivo, nel complesso, è fornire una panoramica sulla modellazione del problema e sulla natura della sfida di creare un solver per Sokoban, rimarcando il modo in cui essa può essere ricondotta a tecniche domain-agnostic di ricerca su un grafo degli stati, ma anche illustrando alcune delle possibili ottimizzazioni domain-specific. Proverò inoltre a documentare in maniera frugale, evitando di perdere eccessivamente di generalità, il lavoro concreto di implementazione delle componenti del solver. Per ulteriori informazioni, si visiti il repo GitHub del progetto.

Concludo l'introduzione con un invito a temperare le proprie aspettative: risolvere Sokoban per un'IA, se non l'avessi già ricordato a sufficienza, non è una passeggiata al parco, e una esplorazione trasversale - e individuale - come quella che seguirà non è sufficiente per giungere a risultati solidi. Basti pensare che a Rolling Stones, il più celebre e performante Sokoban solver esistente, creato dall'University of Alberta, è stato necessario più di un decennio di sviluppo per riuscire a risolvere tutti i 90 puzzle del set standard di Sokoban. È interessante notare che un essere umano, data una quantità sufficiente di Red Bull e disdegno per la propria esistenza, potrebbe riuscirci in una giornata sola. Evidentemente l'IA non è ancora pronta a soppiantarci del tutto, almeno fin quando non imparerà a bere Red Bull e odiare se stessa. Ad ogni modo, buona lettura e buona fortuna con la mia letale verbosità cagionata dalla noia del lockdown e dalle mie passate velleità letterarie.

## 2 – Sokoban solver come agente intelligente

Questo capitolo ha lo scopo di formalizzare le caratteristiche di SokoBOT come agente intelligente, partendo dalla sua rappresentazione PEAS (Performance, Environment, Actuators, Sensors). In altre parole, sarà mostrato che un solver per Sokoban, pur operando nel contesto piuttosto astratto di un livello di gioco, può essere visto come un sistema che incamera una sequenza di **percezioni** su un **ambiente** e opera su di esso con una sequenza di **azioni** definite da una **funzione agente** in relazione a percezioni e azioni pregresse. L'operato dell'agente è guidato da una misura di **prestazione** che valuta il successo delle azioni compiute in base a criteri che saranno ampiamente discussi in seguito.

SokoBOT sarà un **agente basato su obiettivi** con lo scopo di trovare una sequenza di azioni di Sokoban che risolva un livello dato in input. Esso avrà contezza dello stato in cui si trovano se stesso e l'ambiente, nonché dell'effetto delle proprie azioni. Avrà memoria degli stati passati e compierà le proprie scelte, e dunque azioni, con l'unica intenzione di raggiungere uno stato riconosciuto come obiettivo. Le caratteristiche di uno stato obiettivo sono evidenti fin dal principio: SokoBOT tenderà verso uno stato che rappresenti una situazione di risoluzione del puzzle su cui sta operando. Questa tendenza a rappresentare la desiderabilità di uno stato in termini binari può, però, diventare più sfumata nel caso in cui il **programma agente** di SokoBOT basasse le sue azioni su criteri euristici che forniscano dei valori di utilità agli stati che ne rappresentino la desiderabilità, dove per stato *desiderabile* si intende *vicino alla soluzione*. In questo scenario, SokoBOT potrebbe essere interpretato come un ibrido tra un agente basato sugli obiettivi e un agente basato sull'utilità. Nonostante ciò, la natura stessa della "missione" di SokoBOT rende chiara la sua anima da *goal-based agent*: utilità degli stati o meno, esso verrà sempre incaricato di attivarsi, trovare una soluzione ad un puzzle e fermarsi.

### 2.1 – Ambiente e Sensori

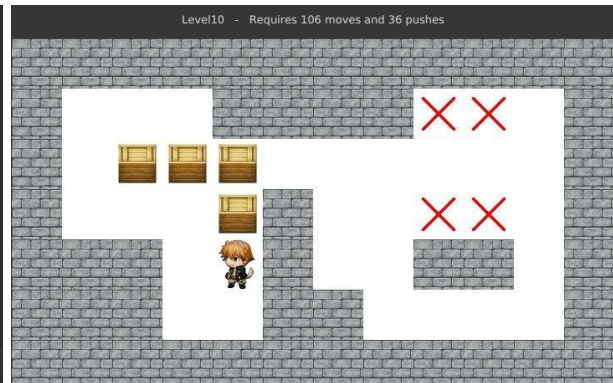
L'ambiente in cui SokoBOT opera è il *maze* contenente il puzzle da risolvere o, più concretamente, una sua rappresentazione matriciale. Dato un qualsiasi livello di Sokoban, esso può essere visto come una griglia di caselle di gioco simile ad una scacchiera, rappresentato da una matrice avente alla posizione  $(i,j)$  un valore che identifica il contenuto della casella di gioco di riga  $i$  e colonna  $j$ . I valori possibili sono i seguenti:

- **E**: casella vuota, su cui è possibile spostare Sokoban o una cassa.
- **S**: casella occupata da Sokoban.
- **B**: casella occupata da una cassa.
- **W**: casella occupata da un muro; essa non è attraversabile e non cambierà mai valore.

Alcune delle caselle che nella rappresentazione iniziale del livello hanno valore E, S o B possiederanno un'etichetta aggiuntiva “goal”, che permette di individuare le caselle obiettivo, ossia quelle in cui bisognerà spingere le casse. Possiamo immaginarla come un valore booleano, che viene registrato al momento della prima rappresentazione dell'ambiente di gioco e non è soggetto a cambiamenti nel tempo: le caselle obiettivo, difatti, possono cambiare contenuto ma non smetteranno mai di essere caselle obiettivo. Nell'implementazione concreta di SokoBOT, l'ambiente è codificato in un file .json contenente, come è possibile immaginare, una matrice di caratteri in cui, alla posizione (i,j), figura la lettera maiuscola corrispondente al contenuto iniziale della casella di gioco alla riga i e colonna j.

```
{
  "content" : [
    [ "W", "W", "W", "W", "W", "W", "W", "W", "W", "W", "W", "W" ],
    [ "W", "E", "E", "E", "W", "W", "W", "W", "G", "G", "E", "W" ],
    [ "W", "E", "B", "B", "B", "E", "E", "E", "E", "E", "E", "W" ],
    [ "W", "E", "E", "E", "B", "W", "E", "E", "G", "G", "E", "W" ],
    [ "W", "W", "W", "E", "S", "W", "E", "E", "W", "W", "E", "W" ],
    [ "W", "W", "W", "E", "E", "W", "W", "E", "E", "E", "E", "W" ],
    [ "W", "W", "W", "W", "W", "W", "W", "W", "W", "W", "W", "W" ]
  ],
  "bestSolution" : 106,
  "minPushes" : 36
}
```

Codifica .json di un livello...



...e lo spazio di gioco corrispondente

La lettura dei file .json relativi ai livelli di gioco rappresenta, per SokoBOT, l'unico input richiesto per formare la conoscenza ambientale necessaria per risolvere il puzzle, e costituisce dunque l'unica sua istanza di “**sensore**”. Parliamo ovviamente di un sensore software, che ha bisogno di entrare in attività una sola volta, al momento di percepire la configurazione matriciale iniziale del livello da risolvere. Da quel momento in avanti, l'agente non ha bisogno di ulteriori sequenze sensoriali. In ogni momento esso può inferire lo stato dell'ambiente sulla base delle azioni compiute su di esso a partire dalla prima percezione, operando in un ambiente **completamente osservabile** e **deterministico**. Questo schema percettivo *one-off* del programma agente di SokoBOT è reso possibile dal fatto che esso può **derivare** stati di gioco successivi al primo, completamente conscio sia delle caratteristiche del *maze*, sia delle conseguenze delle proprie azioni. In altre parole, le rappresentazioni degli stati successivi al primo possono essere sempre calcolati accuratamente, e non è utile introdurre nel modello dell'agente ulteriori percezioni. Del resto, la missione operativa del risolutore è meglio riassunta da:

“Ecco un livello. Come si risolve?”

piuttosto che da:

“Ecco un livello. Come si risolve? Ecco un livello leggermente diverso da quello di prima. Come si risolve? Ecco un altro livello ancora leggermente diverso. Come si risolve? ...”.

Continuando con la classificazione dell'ambiente, i *maze* di Sokoban costituiscono un ambiente **semi-dinamico**: gli unici cambiamenti ad esso scaturiscono dall'intervento dell'agente risolutore, e a nessun poltergeist maligno è concesso di modificare la posizione degli elementi mentre SokoBOT è distratto, ma il passare del tempo incide sulla valutazione delle sue prestazioni. È un

ambiente **discreto** perché il numero di mosse ad ogni passo è limitato, e le possibili configurazioni delle caselle sono finite, pur tendendo spesso ad un numero altissimo.

Lo svolgimento innatamente singleplayer di Sokoban porta in genere a modellare un agente risolutore che operi in un ambiente **ad agente singolo**, e SokoBOT non fa eccezione. È interessante notare, però, che in letteratura ci sono stati vari tentativi, anche di discreto successo, di costruire risolutori modellati per agire su un ambiente multi-agente, interpretando le casse come agenti indipendenti l'uno dall'altro che “usano” Sokoban come strumento per raggiungere la loro meta, attirandolo a sé e dettandogli le spinte più convenienti.

## 2.2 – Attuatori e misure di Prestazione

Discorrere di *attuatori* nel contesto di un software di puzzle solving è certamente un formalismo. L'agente crea una rappresentazione degli stati dell'ambiente in forma di strutture dati ad hoc, e il kernel del sistema operativo che lo ospita gli concede – o meglio, per essere pedanti, concede alla JVM - il sommo potere di modificarle quandunque lo ritenga necessario per generare nuovi stati come conseguenza delle azioni scelte. L'influenza di SokoBOT sull'ambiente di gioco non deve fronteggiare particolari “attriti” e non deve mediare con altre componenti dell'ambiente. Deliberata una mossa di gioco, essa può essere semplicemente materializzata mandando in esecuzione i frammenti di codice che calcolano un nuovo stato e lo conservano in memoria.

Passiamo a qualcosa di decisamente più interessante. L'ovvia, minima e necessaria misura di prestazione di un agente risolutore di Sokoban è **risolvere Sokoban**. Scioccante, vero? Lo scopo dell'agente è produrre una sequenza di azioni che costituisca una soluzione per il puzzle in input, e se non riesce a fare ciò, nessun'altra metrica possibile ha importanza. Questo discorso va però commisurato alla difficoltà del problema che l'agente maneggia, ed essa dipende dalla taglia e dalla complessità del livello. Esistono strategie che assicurano la risoluzione di qualsiasi livello possibile, ma non necessariamente esse hanno costi sostenibili in termini di tempo di computazione e spazio di archiviazione.

Le altre misure di prestazione sono il **tempo di computazione**, la **memoria utilizzata** e il **numero di step** richiesti dalla soluzione generata per portare l'ambiente in uno stato vincente. Le tre misure citate vanno, idealmente, minimizzate, ma per la memoria utilizzata c'è da fare una precisazione. Lo scenario ideale non è strettamente usare *meno* memoria possibile, ma piuttosto di usarla *meglio* possibile. L'agente dovrebbe usare una quantità di memoria abbastanza bassa da non ritrovarsi privato delle risorse necessarie per restare in funzione, ma anche abbastanza alta da non dover scendere a compromessi computazionali non necessari che influenzino altre metriche di prestazione. In altre parole, vale il vecchio mantra: la memoria non usata è memoria sprecata.

Avere un tempo di computazione più basso possibile ha un'importanza piuttosto intuitiva, ma la definizione di *basso* può variare. Se intendiamo il risolutore come un'entità astratta che fornisce

una risposta a un enigma matematico, possiamo permetterci di aspettare diverse ore, perfino giorni o settimane, sebbene un tempo più breve possibile resti comunque decisamente preferibile. Se, invece, prendiamo sul serio il proposito di usare la ricerca su Sokoban come modello per problemi di robotica, dobbiamo considerare che non trarremmo particolare giovamento dal costruire un androide che, incaricato di spostare oggetti in determinati posti, si metta in moto una settimana dopo aver ricevuto il nostro ordine.

Infine, vale la pena di notare che esistono due modi diversi di intendere il *numero di step* richiesti da una soluzione: contare il numero totale di movimenti effettuati da Sokoban, oppure contare solo il numero di spinte alle casse. La questione è particolarmente rilevante perché adottare l'uno o l'altro come criterio principale porta la ricerca di una soluzione su strade sensibilmente diverse. Soluzioni che ottimizzano il numero di movimenti non faranno necessariamente lo stesso per il numero di spinte. Avremo modo di dettagliare abbondantemente questa distinzione più avanti.



## 3 – Sokoban come problema di ricerca

Come già accennato, il compito di SokoBOT non è triviale. Nella maggior parte dei casi, un livello non può essere risolto semplicemente considerando le casse una per una e spingendole verso una casella obiettivo. Anche perché, del resto, non sarebbe affatto divertente per un essere umano risolvere un livello del genere. È necessario solitamente *sbrogliare una matassa* di ostacoli e strettoie, prima di poter vedere chiaramente il percorso verso la vittoria. Per fare ciò, noi esseri umani possiamo impiegare una strategia d'insieme, valutare in maniera organica e flessibile la disposizione delle nostre risorse, e procedere immaginando il vantaggio o lo svantaggio delle varie strategie ideate dalle nostre sinapsi. L'obiettivo di questo capitolo è di introdurre gli ingredienti necessari per definire il problema di risolvere Sokoban con un agente risolutore che, essendo stato partorito da un wafer di silicone, non gode dei lussi cognitivi di cui sopra.

### 3.1 – Stati e transizioni

Per risolvere Sokoban computazionalmente, occorre vedere una partita come una serie di **transizioni** tra **stati**, determinate dal compimento delle azioni disponibili. L'insieme di tutti gli stati che è possibile raggiungere durante una partita è lo **spazio degli stati**, ed è il teatro dalle quinte immensamente larghe in cui un agente risolutore mette in scena la propria ricerca di una soluzione. Entriamo ora nello specifico dei singoli stati e delle transizioni tra essi. Possiamo dare due definizioni diverse di stato, che risulteranno familiari se si è prestata attenzione a quanto detto sui modi di *contare gli step* nel Par. 2.2.

Possiamo definire uno stato nello spazio di ricerca come una configurazione unica degli elementi semoventi dell'ambiente, ossia ognuna delle casse e Sokoban. Questa definizione implica che muovere Sokoban in qualsiasi modo, che sia spinta una cassa o meno, determina la genesi di un nuovo stato. Adottare questa definizione di stato porta a fissare il numero massimo di nuovi stati generabili in ogni momento a esattamente quattro, uno per ognuna delle quattro direzioni in cui Sokoban può muoversi.

Alternativamente, uno stato può essere definito come una configurazione unica delle casse. In questo caso, ignoriamo il libero girovagare di Sokoban e decidiamo che una configurazione sarà giudicata un nuovo stato soltanto quando una cassa sarà spinta e cambierà posizione. Due stati A e B possono essere considerati equivalenti se le casse sono nella stessa posizione, e se Sokoban può raggiungere, senza spingere casse, la posizione che occupa nello stato B partendo dalla posizione che occupa nello stato A e viceversa. Questa scelta aumenta il numero massimo di nuovi stati generabili a partire da un dato stato a quattro moltiplicato per il numero delle casse nel livello, poiché, al più, in un dato momento Sokoban può muovere ognuna delle casse in quattro direzioni diverse. Sebbene apparentemente si possa riscontrare una più rapida proliferazione degli stati, questa seconda definizione è la più utile, efficace e comune. Per darmi un tono, ho deciso di battezzare la prima definizione **move formulation** e la seconda **box formulation**.

## 3.2 – Formalizzazione del problema

Possiamo procedere a fornire, per ognuna delle formulazioni della definizione di stato, una descrizione formale completa del problema.

### Move formulation

- **Stati:** l'insieme di tutte le possibili disposizioni discrete di Sokoban e delle casse all'interno del livello di gioco.
- **Stato iniziale:** il livello di gioco nel suo stato originale, ossia con casse e Sokoban disposti esattamente come prescritto dal file in input.
- **Azioni:** è possibile muovere Sokoban in ognuna delle quattro caselle adiacenti a quella in cui si trova nello stato corrente; notare che un'azione non ha nessun effetto se si sceglie di muovere Sokoban in una casella occupata da un muro o da una cassa che non può essere spostata. Formalmente, Azioni = {move\_up, move\_down, move\_left, move\_right}.
- **Modello di transizione:** una funzione che, dato uno stato di gioco  $s$  e una mossa di Sokoban  $m$ , associa alla coppia  $(s, m)$  un nuovo stato  $s'$  che rappresenti la configurazione del livello di gioco raggiunta applicando la mossa  $m$  allo stato  $s$ .
- **Test obiettivo:** si controlla se nello stato corrente tutte le casse sono su una casella obiettivo.
- **Costo di cammino:** il costo di una sequenza di azioni è pari alla cardinalità della sequenza stessa, ossia al numero di mosse compiute.

### Box formulation

- **Stati:** l'insieme di tutte le possibili disposizioni discrete delle casse all'interno del livello di gioco.
- **Stato iniziale:** il livello di gioco nel suo stato originale, ossia con casse e Sokoban disposti esattamente come prescritto dal file in input.
- **Azioni:** è possibile spingere le casse verso ognuna delle quattro caselle adiacenti a quella occupata da tale cassa nello stato corrente; notare che un'azione non ha nessun effetto se si sceglie di spostare la cassa in una casella occupata da un muro o da un'altra cassa. Formalmente:  $\text{foreach cassa } c \rightarrow \text{Azioni}_c = \{c.\text{push\_up}, c.\text{push\_down}, c.\text{push\_left}, c.\text{push\_right}\}$
- **Modello di transizione:** una funzione che, dato uno stato di gioco  $s$ , una cassa  $c$  e una mossa di spinta  $m$ , associa alla tupla  $(s, c, m)$  un nuovo stato  $s'$  che rappresenti la configurazione del livello di gioco raggiunta a partire dallo stato  $s$  applicando la mossa  $m$  alla cassa  $c$ .
- **Test obiettivo:** si controlla se nello stato corrente tutte le casse sono su una casella obiettivo.
- **Costo di cammino:** il costo di una sequenza di azioni è pari alla cardinalità della sequenza stessa, ossia al numero di spinte compiute.

Si noti che un agente risolutore che adotti la box formulation produrrà, almeno in linea teorica, una sequenza di spinte alle casse e non una sequenza di mosse di Sokoban. L'implementazione della box formulation prodotta per il programma agente di SokoBOT, però, si curerà di “unire i puntini” tra una spinta e l'altra, fornendo così la possibilità di avere la soluzione nella forma più pratica possibile, cioè come sequenza di azioni di Sokoban.

### 3.3 – Schemi di espansione dei nodi nell'albero di ricerca

Il compito di cercare uno stato desiderabile in un sistema di stati e transizioni si configura convenzionalmente come una ricerca su un albero implicito, di cui possediamo soltanto il nodo radice che, nel caso di Sokoban, rappresenta lo stato di partenza del puzzle. Durante l'esecuzione di un algoritmo di ricerca, i nodi visitati, a partire dalla radice, sono **espansi**: vengono “scoperti” e aggiunti all'albero tutti i nodi adiacenti a un nodo appena visitato. In termini di ricerca tra stati, questo ci permetterà di esplorare tutti gli stati a una sola transizione di distanza dallo stato in cui ci troviamo. Logica e ordine di tale esplorazione dipendono dall'algoritmo scelto, ma l'idea è sempre quella di procedere, espansione dopo espansione, ad esplorare nodi generati fin quando non si giunga ad esplorare un nodo che rappresenti uno stato obiettivo. Uno stesso stato nello spazio può essere raggiunto con sequenze di azioni diverse tra di loro, e il risultato è che durante la ricerca costruiremo un **albero con cammini ciclici**, che presenta cioè cammini con più copie dello stesso nodo. Come è facile immaginare, questa caratteristica è un elemento problematico, che può far esplodere la complessità di tempo e di spazio della ricerca, se non affrontato. Breve nota: a seguire, conterò sul fatto che il lettore abbia conoscenze pregresse sulla teoria dei grafi e sugli algoritmi di ricerca su grafi.

Torniamo sulle definizioni di stato, stavolta consci del fatto che rappresenteremo gli stati come nodi di un grafo di ricerca. Definire gli stati, e dunque i nodi di ricerca generati, in base alle spinte alle casse porta il **branching factor** del grafo a  $4 \times b$ , dove  $b$  è il numero delle casse presenti nel livello. Ma definire gli stati in base alle mosse aumenta la **profondità massima** della ricerca, perché, generalmente, le soluzioni prevedono un numero ottimale di mosse ben superiore al numero ottimale di spinte. La questione è rilevante perché la profondità del percorso di ricerca tende, negli algoritmi di ricerca su grafi, ad avere un peso computazionale maggiore rispetto al branching factor. Ecco spiegato il motivo per cui è generalmente preferibile adottare la box formulation rispetto alla move formulation. Espandere i nodi in base alle mosse porta a generare un numero molto più alto di stati, la maggior parte dei quali non sono particolarmente utili nella ricerca della soluzione: valutare anche la possibilità che Sokoban se ne vada a zonzo per il livello senza curarsi delle casse che è pagato per spingere può essere un'idea gradevole per i procrastinatori, ma di certo non avanza la quest dell'agente risolutore.

Nell'implementazione concreta di SokoBOT sarà presente la possibilità di utilizzare entrambi gli schemi di espansione dei nodi per osservarne l'efficacia. Al tempo della stesura di questo paragrafo, sono perlopiù entrambi pronti e testabili ed è già evidente che con l'espansione in base alle mosse si incorre in grandi difficoltà già in puzzle di complessità moderata, con l'esplosione

incontrollabile e indesiderabile del numero degli stati, specialmente in livelli “aperti”, con tanto spazio a disposizione per i peregrinaggi infruttuosi di Sokoban.

Spezzando una lancia in favore della move formulation, c'è da dire che effettuare la ricerca della soluzione espandendo i nodi in base alle possibili spinte delle casse non può garantire – assumendo, per il momento, di utilizzare un ipotetico algoritmo perfetto che trova sempre una soluzione ad ogni puzzle – l'ottimalità della soluzione in termini di mosse totali, ma solo in termini di spinte totali. Peraltro, implementare uno schema di espansione dei nodi che asseconi la move formulation risulta molto più semplice e naturale: se ci troviamo in uno stato, per espanderlo e produrre la sua progenie non dobbiamo fare altro che eseguire tutti e quattro i movimenti disponibili e immagazzinare gli stati così prodotti in nuovi nodi. Terremo conto, ovviamente, del fatto che potrebbero esserci ostacoli intorno a Sokoban, e dunque alcune delle mosse potrebbero non essere effettuabili.

Passando alla box formulation, la situazione è più complessa sia concettualmente che computazionalmente. Per generare gli stati adiacenti a quello corrente, dovremo provare tutte le spinte a una cassa possibili. Ciò significa, innanzitutto, individuare le caselle che contengono una cassa, poi procedere, cassa per cassa, nel seguente modo:

1. Si individuano le caselle libere adiacenti alla cassa e si verifica che dal lato opposto ci sia un'altra casella libera che possa ospitare la cassa. Le caselle che soddisfano tale condizione saranno quelle presso cui Sokoban potrà muoversi per procedere con la spinta della cassa. Si noti che ogni cassa potrebbe essere disponibile per più spinte diverse, fino ad un massimo di quattro.
2. Si verifica che esista un percorso nel maze tra Sokoban e le caselle “di spinta” individuate nel passo precedente. Questo percorso dovrà essere praticabile senza la necessità di spostare altre casse. Nel caso non fosse possibile per Sokoban raggiungere una delle caselle “di spinta”, una particolare spinta non sarà effettuabile, e si procederà a valutare la prossima casella.
3. Se invece quel percorso esiste, lo si determina passo per passo, curandosi che sia il più breve percorso possibile tra Sokoban e la casella da raggiungere.
4. Infine, si muove Sokoban lungo il percorso individuato e si esegue la spinta, generando un nuovo stato secondo la box formulation.

I passi 2 e 3, pur concettualmente distinti, sono stati implementati in congiunzione in SokoBOT, immaginando le caselle di gioco come un – ulteriore – grafo in cui caselle vuote adiacenti siano unite da archi e lanciando su di esso una **ricerca BFS**. La ricerca si arresta con un successo quando la casella di spinta che cercavamo viene trovata: la natura della ricerca in ampiezza ci garantisce che siamo arrivati al nostro obiettivo con il numero minimo di passi. Ovviamente, se arrivassimo ad aver esplorato tutte le caselle vuote senza trovare l'obiettivo, concluderemmo che la casella cercata non era raggiungibile.

Per informazioni più specifiche sull'implementazione degli schemi di espansione in SokoBOT, fare riferimento al codice nel repo GitHub. Ho provato a commentarlo copiosamente, proprio per poter evitare di scendere in dettagli implementativi in questa sede.

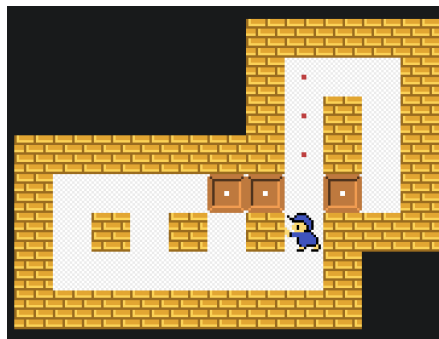
## 4 – Enter SokoBOT

Introdotta il problema di ricerca in generale, passiamo a stabilire come un solver come SokoBOT può procedere per risolverlo. La trattazione dell'argomento sarà incentrata sui metodi di risoluzione che sono – o saranno – implementati concretamente in SokoBOT, ma non mancheranno riferimenti ad altre strategie e ottimizzazioni possibili e documentate. Si noti che molte specifiche riguardanti concetti generali di ricerca informata e non informata saranno date per scontate. Non è particolarmente d'interesse, in questo contesto, fornire una panoramica dettagliata su strategie trasversali di graph search, che saranno toccate *en passant*.

### 4.1 – Elementi di difficoltà

Come già accennato in varie occasioni, lo spazio degli stati su cui operare la ricerca ha una taglia importante. Assumendo di espandere i nodi durante la ricerca per spinte e non per mosse, sia il **fattore di branching** che la **profondità della soluzione** all'interno del grafo di ricerca tendono ad assumere valori molto alti. Nei 90 puzzle del set standard di Sokoban il fattore di branching più alto incontrato ammonta a 136, con una media pari a 12, mentre la profondità della soluzione varia tra 97 e 674. Inoltre, l'interezza dello spazio di ricerca conta, in media,  $10^{18}$  stati distinti, rendendo chiara l'esigenza di imboccare quanto prima una direzione che indirizzi la ricerca su percorsi che abbiano probabilità migliori di terminare con la soluzione. Una ricerca completa in questo spazio degli stati, visitando cioè tutti i nodi per ciascun livello di profondità fino al raggiungimento della soluzione, può risultare esosa in termini di tempo e/o di spazio.

Oltre all'intrattabilità meramente numerica, la natura stessa del gioco introduce un ulteriore ostacolo: la presenza di **deadlock**, configurazioni di stallo in cui risulta impossibile, continuando la partita, arrivare a vincerla. Il dramma però non consiste nella loro esistenza, quanto nella difficoltà di rilevarli e interrompere la ricerca in caso se ne incontri uno. A volte, è difficile perfino per un essere umano accorgersi di aver disposto le casse in una maniera che annulla irreversibilmente le chance di vittoria.



*Esempio di una situazione di deadlock*

Aver creato una configurazione di deadlock non significa aver creato una condizione in cui non sia più possibile muovere Sokoban o le casse. Con un approccio puramente algoritmico, e senza l'ausilio di meccanismi ad hoc di *deadlock detection*, non si ha nessuna possibilità di accorgersi del vicolo cieco e tagliare il ramo corrispondente della ricerca. Al contrario, si continuerà a valutare movimenti e spinte che non possono più portare alla soluzione. Esistono, per onor del vero, casi in cui, almeno se si classificano gli stati per box formulation, avviene automaticamente una sorta di deadlock detection. Nel caso in cui spingessimo le casse in degli angoli del livello, esse non potranno più essere spostate, e dunque quel percorso di ricerca terminerà. Ovviamente si tratta di un fringe case: i deadlock più comuni durante una partita sono molto più sottili e difficili da rilevare.

## 4.2 – Rappresentazione e trasposizione degli stati

Durante la ricerca, SokoBOT usa oggetti *Node* dedicati alla rappresentazione di singoli stati nello spazio di ricerca. Per maggiore praticità nel processing dei dati essi comprendono, oltre alla matrice che tiene traccia del contenuto di ognuna delle caselle di gioco – incapsulata in un oggetto *GameBoard* –, una pletora di informazioni sia sullo stato rappresentato che sul nodo nell'albero di ricerca, ad esempio un riferimento diretto alle caselle che contengono le casse, oppure il livello di profondità nell'albero in cui il nodo è stato trovato. La classe *Node* non è un POJO inerte, ma implementa anche la logica delle operazioni sui nodi che gli algoritmi di risoluzione dovranno effettuare, tra cui l'espansione di un nodo.

In maniera del tutto trasversale alla strategia di ricerca adottata, è necessario effettuare una prima, essenziale forma di pruning dello spazio degli stati, ossia impedire la nascita di **cammini ciclici** nell'albero di ricerca. In altre parole, è necessario evitare che l'espansione di un nodo porti a generare nodi che rappresentino stati già esaminati, portando lo spazio di ricerca a crescere in maniera incontrollabilmente superiore rispetto all'effettivo spazio degli stati. Senza questa basilare accortezza, anche risolvere livelli estremamente piccoli diventa un piccolo miracolo computazionale. Fortunatamente, è relativamente facile ottenere il risultato sperato implementando una struttura dati che memorizzi delle **trasposizioni** degli stati toccati, ossia delle rappresentazioni sintetiche e – se tutto va bene – univoche di tali stati. Ogni volta che un nuovo nodo viene generato durante un'espansione, si controlla se la sua trasposizione è già presente: in quel caso, possiamo potare quel ramo. Nel caso in cui non fosse presente, invece, possiamo procedere esplorando anche in quella direzione, certi del fatto che stiamo rappresentando uno stato inedito. Ovviamente, ad ogni nuova “accettazione” di un nodo espanso, dobbiamo inserirne la trasposizione nella struttura.

Le trasposizioni in SokoBOT sono implementate trasformando la board di gioco in un array di byte e producendone il valore hash MD5, memorizzato in una variabile long e inserito nella struttura

delle trasposizioni. L'unica componente ad essere coinvolta nell'hashing è la matrice di celle che rappresentano uno stato della board di gioco, mentre i campi "di servizio" sono ignorati per evitare comportamenti indesiderati e sprechi di spazio. Nodi che possiedono board di gioco in configurazione identica avranno una trasposizione identica: basterà dunque consultare la lista di tali trasposizioni per verificare se abbiamo già incontrato un stato. L'algoritmo di hashing MD5, pur non essendo crittograficamente sicuro da decenni, ha una resistenza alle collisioni che lo rende più che sufficiente per evitare che due stati diversi siano trasposti con il medesimo valore hash.

Si noti che, nel caso si stiano espandendo i nodi in funzione delle spinte alle casse, non sarà sufficiente trasporre la board di gioco nell'esatto stato in cui la si trova. Agendo in questo modo, infatti, trasporremmo gli stati tenendo conto della posizione di Sokoban. Due trasposizioni, anche presentando le casse nella stessa posizione, sarebbero comunque ritenute uniche nel caso Sokoban figurasse in posizioni diverse, determinando l'esplorazione di stati che, per la box formulation, non sarebbero nuovi stati. Per rimediare, possiamo omettere Sokoban dalla rappresentazione prima di effettuare l'hashing. Dobbiamo però accertarci che, nel confrontare le trasposizioni di due stati A e B con la stessa configurazione delle casse, questi siano ritenuti identici solo se esiste un percorso tra la posizione di Sokoban in A e la posizione di Sokoban in B e viceversa. Il modo più semplice per assicurare tale condizione è codificare le trasposizioni in modo da prendere in considerazione la configurazione della board priva di Sokoban, con l'aggiunta della lista ordinata delle caselle raggiungibili da Sokoban in tale configurazione. Trasposizioni congruenti in tale codifica assicurano di rappresentare solo stati congruenti secondo la box formulation.

## 4.3 – Algoritmi di ricerca non informata

Risolvere Sokoban lanciando una ricerca cieca senza sfruttare conoscenza aggiuntiva per orientare la navigazione nell'albero di ricerca potrebbe sembrare futile. In effetti, gli effort di maggiore successo sono storicamente basati su algoritmi di ricerca informata. Gli algoritmi di ricerca non informata hanno però il merito intrinseco di non aver bisogno di una stima euristica per produrre risultati. Sebbene appaia sciocco non ottimizzare la ricerca con criteri euristici, laddove fossero disponibili, occorre ricordare che le valutazioni euristiche non piovono sempre dal cielo. La loro elaborazione potrebbe essere computazionalmente intensa, e potrebbe richiedere l'uso di strutture dati aggiuntive che peggiorano l'occupazione di memoria. Ad esempio, proprio la stima euristica "corretta" della bontà di uno stato in Sokoban, che descriverò in seguito, non è delle più semplici da calcolare, e domina il tempo d'esecuzione degli algoritmi di ricerca informata. La ricerca non informata aumenta la fetta disponibile del budget computazionale tagliando i costi di calcolo al momento di espandere i nodi. Gli algoritmi implementati in SokoBOT sono **BFS** e **Iterative Deepening DFS**.

Il primo è una implementazione semplicissima e assolutamente convenzionale di una ricerca in ampiezza, con le consuete caratteristiche: ottimalità nel caso in cui il costo di cammino cresca insieme alla profondità dei nodi, completezza, e un investimento potenzialmente inaccettabile di

memoria. Abbinato a una struttura di trasposizioni efficace, esso è già sufficiente per trovare la soluzione ottima a puzzle di taglia ridotta, e mi è stato dunque utile per testare l'impalcatura software del solver prima di iniziare a perfezionare le meccaniche di risoluzione.

La ricerca in profondità è stata implementata con una meccanica di approfondimento iterativo. La DFS *vanilla*, infatti, è ben poco adatta a fare ricerche su un albero implicito di taglia potenzialmente enorme: rischierebbe di approfondire un solo percorso per moltissimo tempo prima di fare backtracking. L'iterative deepening è stato implementato in una variante che produce un uso intelligente della memoria di sistema. Per ognuna delle profondità di arresto, tutti i nodi "di frontiera" su cui la ricerca si è arrestata vengono salvati in una struttura **cache**. In questo modo, l'iterazione successiva approfondita non dovrà ripartire dalla radice dell'albero: potrà usare la frontiera salvata come nodo di partenza, risparmiando il tempo di computazione che sarebbe necessario per giungere di nuovo a quel livello. Questo procedimento viene iterato solo fino a quando è disponibile memoria sufficiente. Nel momento in cui si constata di non avere più memoria, si passa ad una ricerca ad approfondimento iterativo classica, usando però come punto di partenza non la radice dell'albero, ma l'ultima frontiera salvata.

Ciò permette di risparmiare tempo di computazione nei livelli meno profondi dell'albero, rendendo la risposta di SokoBOT su livelli di taglia moderata pressoché istantanea. A livello asintotico, però, non c'è quasi nessuna differenza. Il tempo di computazione cresce esponenzialmente al crescere della profondità. Ci possiamo perciò aspettare che il tempo necessario per arrivare dal livello di profondità 1 al livello  $n-1$  sia abbondantemente minore rispetto al tempo necessario per computare il passaggio dal livello  $n-1$  al livello  $n$ , per valori di  $n$  molto alti. A conti fatti, salvare la frontiera è una miglioria che può portare con sé la comodità di ridurre a pochi secondi il tempo di una ricerca che potrebbe durare qualche minuto. Ma, su una ricerca dalla durata prevista di una settimana, riuscirebbe probabilmente a risparmiare solo qualche ora di calcolo.

È interessante notare che, se si fosse implementata una ricerca a costo uniforme, la natura dell'espansione di nodi nello spazio di ricerca di Sokoban l'avrebbe fatta collassare in una ricerca in ampiezza. Nel nostro problema di ricerca, ogni volta che si espande un nodo la sua prole rappresenta tutti gli stati raggiungibili con una singola mossa o spinta. L'ovvia conseguenza è che ogni nodo nel *batch* così prodotto avrà costo di cammino 1 più il costo del genitore: tutti i nodi ad un certo livello di profondità avranno lo stesso costo di cammino. Una ricerca a costo uniforme non farebbe altro che esplorare tutti i livelli di profondità uno dopo l'altro, visto che non troverà mai un nodo in un livello che abbia costo inferiore ai propri compagni di livello, e allo stesso modo non troverà mai un nodo al livello  $n+1$  che abbia costo inferiore rispetto a qualche nodo al livello  $n$ . In tal modo, non "sfalserà" mai il processo di ricerca e procederà esattamente come se da grande avesse sempre sognato di fare la BFS.



## 4.4 – Euristiche di stima e algoritmi di ricerca informata

Un modo piuttosto comune di migliorare le prestazioni di un agente risolutore impegnato in imprese complesse come quella con cui si sta scontrando SokoBOT è *informare* la ricerca individuando un criterio euristico per stimare la bontà di uno stato. Esso sarà distillato matematicamente in una **funzione euristica** che ci fornisca, dato un qualsiasi stato, una stima più o meno accurata della sua distanza dalla soluzione.

Consideriamo di trovarci all'interno di un maze di gioco, e di valutare i passi compiuti in termini di spinte alle casse (box formulation). Quanto siamo distanti dalla soluzione? La risposta sarà, in qualche modo, una funzione della distanza delle casse dagli obiettivi o, per essere più precisi, della **Manhattan distance** delle casse dagli obiettivi, ossia la distanza tra due punti in un piano in cui, esattamente come negli isolati newyorkesi, ci si può muovere solo lungo una griglia. Ora, dobbiamo formalizzare una funzione euristica **ammissibile** – che fornisca, cioè, un lower bound della soluzione – usando le distanze tra le casse e gli obiettivi.

Possiamo vedere casse e obiettivi come le due partizioni di un grafo bipartito, e creare un **matching** che unisca ogni cassa all'obiettivo ad essa più vicino, dopodiché sommare le Manhattan distance calcolate tra i membri delle coppie così formate. La funzione appena descritta è sicuramente un'euristica ammissibile, ma è abbastanza facile notare che può sottostimare brutalmente la distanza dalla soluzione: è un lower bound molto largo. Se due o più casse si trovano in prossimità di uno stesso obiettivo, entrambe saranno “matchate” con esso, ma solo una cassa può occupare ciascun obiettivo, e quindi il reale lower bound della soluzione sarà probabilmente molto più alto. Avremmo una sottostima quasi imbarazzante nel caso estremo in cui tante casse fossero vicine ad un solo obiettivo e tutti gli altri obiettivi si trovassero molto distanti.

È dunque chiaro che il matching tra le due partizioni debba essere perfetto: a ogni cassa deve essere associato un obiettivo diverso. La soluzione matematicamente approvata al nostro dilemma è quindi calcolare il **matching perfetto di valore minimo**, ossia quello che minimizza la somma delle Manhattan distance delle coppie scelte. Tale compito è possibile in tempo polinomiale grazie all'**algoritmo ungherese**, anche noto come algoritmo di Kuhn-Munkres. Si tratta di un algoritmo di ottimizzazione matematica che opera su sistemi di vincoli esprimibili in forma matriciale, ma può essere espresso in termini di assegnazioni su grafi bipartiti. Il tempo d'esecuzione è polinomiale, ma non particolarmente veloce –  $O(n^4)$  –, determinando, sulle migliaia di esecuzioni richieste per valutare gli stati toccati da un algoritmo di ricerca informata, un sensibile overhead. Sarà possibile confrontarne le prestazioni in SokoBOT con l'euristica naïf di assegnazione minima greedy descritta nel paragrafo precedente.

Stabilita un'euristica di valutazione – o un array di euristiche da sfruttare in una qualche combinazione – non resta che implementare un algoritmo che possa trarne vantaggio. La scelta è ricaduta su **A\***. Esso, seppure stravolto da decine di ottimizzazioni domain-specific, è alla base del già citato *Rolling Stones* dell'Università di Alberta, ed è generalmente considerato una scommessa sicura. Le due implementazioni di **A\*** presenti in SokoBOT sono abbastanza diverse tra loro. La prima implementa **Simplified Memory Bounded A\***, ossia **A\* vanilla** con l'aggiunta di un meccanismo di **eviction** dei nodi meno promettenti. La seconda implementa **Iterative Deepening**

A\*, che è essenzialmente equivalente a IDDFS, ma usa i valori  $f(n)$  - costo di cammino più stima euristica - dei nodi invece della loro profondità per stabilire il punto di cutoff delle singole iterazioni. Per IDA\* è stato usato lo stesso pattern di mantenimento in memoria delle frontiere già descritto per IDDFS.

## 4.5 – Deadlock detection