

# Federico Giuseppe Rubino

## *Composizione artificiale e creatività computazionale: lo stato dell'arte*

### Introduzione

Negli ultimi 60 anni l'evoluzione dei sistemi computazionali ha portato a esplorare anche il rapporto tra questi ultimi e una loro possibile creatività. Oggi gli sviluppi e gli studi nel campo dell'Intelligenza Artificiale (AI) hanno permesso di generare, interpretare e innovare in molti ambiti della scienza e dell'attività umana. A questa non si sottrae la musica che risulta particolarmente interessante in quanto da un punto di vista scientifico non necessita di una semantica esplicita come altre forme d'arte e offre svariati modi di essere rappresentata matematicamente e computazionalmente<sup>1</sup>. Per questo motivo ricercatori di più branche, ingegneri, psicologi, musicologi ecc... si sono posti l'obiettivo di comprendere se e come si possa parlare di AI creativa. Questo lavoro vuole offrire una sintesi dello stato dell'arte sulla capacità di un sistema computazionale di essere creativo in ambito musicale. Ma prima di analizzare i principali sviluppi è importante ripercorrere alcuni passaggi storici e concettuali condivisi dagli attuali filoni di ricerca per consolidare un linguaggio comune.

Questo elaborato espone le definizioni di creatività per poi esplorare il concetto di creatività computazionale passando brevemente anche per l'evoluzione dell'AI dalla sua nascita all'approdo di ciò che oggi possiamo chiamare intelligenza artificiale creativa.

### 1. Creatività

#### 1.1 Formalizzare la Creatività

Tra gli studiosi più autorevoli nell'ambito della creatività e dell'AI, M.A. Boden afferma che la creatività è una caratteristica dell'intelligenza umana radicata anche nella quotidianità delle azioni, e non soltanto propria di un

<sup>1</sup> G.A. Wiggins, *Creativity, information, and consciousness: The information dynamics of thinking*, in «Physics of Life Reviews», fasc. 34-35, 2020, pp. 1-39:6.

contesto artistico, che coinvolge sia una dimensione cognitiva (generazione di nuove idee, dimensione su cui si concentrano gli attuali modelli di AI) sia una dimensione emotiva, legata ad un preciso contesto culturale e a fattori personali<sup>2</sup>. Questa definizione necessita di una distinzione tra una creatività quotidiana fatta di associazioni di idee, percezioni, autocritica, e una di tipo artistico delimitabile in un determinato contesto come la musica, la letteratura e le arti visive. Per formalizzare una caratteristica così complessa del comportamento umano e non riducibile solo al prodotto finale, la creazione, è utile una sua divisione in parti. A questo proposito Anna Jordanous, nell'ambito degli studi sulla valutazione della creatività, ripropone una suddivisione della stessa, ideata inizialmente da M. Rhodes per studi in contesto didattico, in quattro prospettive, quattro 'P'<sup>3</sup>:

- *Person*: l'agente creativo, che esso sia umano o che sia computazionale
- *Process*: il processo creativo
- *Product*: il prodotto del processo creativo
- *Press*: il contesto in cui la creatività si sviluppa

Delle quattro 'P' il processo creativo è ciò che più ha interessato la ricerca nel campo della creatività computazionale perché per ottenere un comportamento creativo da un computer bisogna sapere descrivere la creatività in termini matematici<sup>4</sup>.

Rimanendo sul piano della creatività artistica essa consiste nel generare qualcosa di nuovo. I.L. Gotz fa notare come però il concetto di creatività sia strettamente legato al concretizzare, ad un processo creativo che porti ad un risultato, e come il concetto di creatività vada separato da quello di intuizione, illuminazione o scoperta sebbene la creatività sia intimamente correlata ad esse<sup>5</sup>.

## 1.2 Categorizzare la creatività

Tra le varie modalità utilizzate dagli studiosi per categorizzare la creatività entro un insieme di attività umane, pare interessante rilevare come essa consista nella rielaborazione e nello sviluppo di conoscenze pregresse. M. Boden a questo proposito differenzia la creatività in<sup>6</sup>:

- *creatività combinatoria*: consiste nel generare nuove idee combinando elementi di conoscenze pregresse.

<sup>2</sup> V. Deary, *Dimensions of Creativity*, edited by M. Boden, MIT Press, Cambridge, MA, 1994, in «Psychological Medicine», fasc. 28, 2, 1998, pp. 485-494.

<sup>3</sup> A. Jordanous, *Four PPP Perspectives on computational creativity in theory and in practice*, in «Connection Science», fasc. 28, 2 2016, pp. 194-216:195.

<sup>4</sup> F. Carnovalini, A. Rodà, *Computational Creativity and Music Generation Systems: An Introduction to the State of the Art*, in «Frontiers in Artificial Intelligence», fasc. 3, 2020, art. 14, p.3.

<sup>5</sup> I.L. Gotz, *On Defining Creativity*, in «The Journal of Aesthetics and Art Criticism», fasc. 39, 3 1981, pp. 297-301:300.

<sup>6</sup> M.A. Boden, *Creativity and artificial intelligence*, in «Artificial Intelligence», fasc. 103, 1998, pp. 347-356:348. Qui e in seguito la traduzione dai testi inglesi è mia.

– *creatività esplorativa*: consiste nel generare nuove idee partendo dall'esplorazione di «spazi concettuali strutturati»

– *creatività trasformativa*: implica la trasformazione di una o più dimensioni dello spazio concettuale, quindi la nuova idea non appartiene allo spazio concettuale originale ma cambia la forma dello stesso.

Importante è proprio l'introduzione da parte di Boden dell'idea dello spazio concettuale, cioè «una particolare struttura per rappresentare informazioni a livello concettuale»<sup>7</sup> o, come definito da F. Carnovalini e A. Rodà sempre in riferimento ai lavori di M. Boden, uno «spazio dove esistono i possibili concetti, alcuni dei quali sono stati esplorati, altri ancora da scoprire»<sup>8</sup>. Ed è in questo spazio concettuale, cioè lo «spazio di tutte le cose che potrebbero essere generate secondo un insieme di regole»<sup>9</sup>, come può essere l'insieme di regole che definiscono la musica, che l'atto creativo artistico prende forma.

### 1.3 Valutare la creatività

Per verificare l'effettiva capacità creativa di una macchina sono necessari dei parametri secondo cui valutare la creatività, a prescindere dall'essere umana o computazionale. Questi parametri devono potersi applicare sia alla creatività umana che a quella computazionale in quanto l'obiettivo di quest'ultima è quello di essere indistinguibile e al pari della creatività umana. Un approccio classico è quello del Test di Turing, che in origine non è pensato per valutare un sistema di creatività computazionale ma può essere facilmente adattato a questa casistica. Funziona in questo modo: si pone davanti ad una giuria un artefatto di creazione umana e uno di creazione computazionale. Se la giuria non capisce quale dei due artefatti è prodotto dalla macchina allora la macchina passa il test. Sorge però una limitazione del test messa in evidenza da C. Ariza e cioè come in realtà questo test, e test simili, non valutino l'effettiva capacità creativa della macchina ma piuttosto la qualità dell'imitazione di un prodotto umano.

Un criterio molto comune è quello introdotto da M. Boden: un prodotto è frutto di un processo creativo solo se è nuovo e ha un valore. A. Jordanous spiega come questo criterio sia applicabile alle quattro P della creatività in cui è presente anche il processo, il che rispetterebbe l'intenzione iniziale, cioè quella di valutare sia il prodotto che il processo<sup>10</sup>. Se si dovesse usare questo criterio è importante specificare per quale tipo di pubblico la creazione deve avere un valore, in quanto il concetto di valore artistico è percepito in maniera soggettiva.

<sup>7</sup> M.A. Boden, *Conceptual Spaces in Milieus of Creativity* cap. 12, a cura di P. Meusburger et al., Springer Netherlands, Dordrecht, 2009, vol. 2, pp. 235-243: 238.

<sup>8</sup> Carnovalini, Rodà, *Computational Creativity and Music Generation Systems*, p. 4.

<sup>9</sup> C. Lamb et al., *Evaluating Computational Creativity: An Interdisciplinary Tutorial*, in «ACM Computing Surveys», fasc. 51, 2 2019, pp. 1-34:4.

<sup>10</sup> *Ibidem*.

Esistono molti altri metodi che si basano su teorie della creatività diverse ma altrettanto valide. Alcuni metodi di valutazione sono stati nel tempo formalizzati e automatizzati, ma non c'è un accordo comune su quale sia il metodo migliore. Inoltre con l'avvento del *Deep Learning* è sorto il problema per cui non è possibile conoscere il processo creativo utilizzato dalla macchina in quanto esso è deciso in autonomia dalla stessa (cfr. par. 2.3).

## 2. Creatività computazionale e sue potenzialità

Una volta provato a definire il concetto di creatività bisogna chiedersi se un computer possa essere creativo. È importante capire che non si può valutare un computer solo in funzione del suo prodotto; infatti il computer potrebbe generare qualcosa di nuovo partendo da concetti non musicali e utilizzando solamente un sistema di calcolo combinatorio non legato a elementi musicali (cfr. par. 2.2). I tentativi di usare la casualità in una produzione creativa non umana sono stati molteplici. In campo musicale ad esempio, anche se con obiettivo ludico, una composizione creata in maniera casuale si può ottenere con il *Musikalisches Würfspiel*, tradotto con 'gioco musicale dei dadi', di W.A. Mozart. Mozart compose 176 battute numerate indipendenti tra loro e definì dei criteri secondo cui con il lancio di due dadi è possibile comporre un minuetto estraendo in successione dalle 176 battute. I criteri sono raccolti nelle due tabelle in Fig. 1. Nelle righe è presente il risultato della somma del lancio dei due dadi. Nelle colonne è presente il numero del lancio, in tutto 16 lanci. L'incrocio tra righe e colonne porta ad una delle 176 battute. La composizione che si ottiene così facendo sarà completamente casuale, tenendo conto delle regole matematiche della probabilità per cui dalla somma del lancio di due dadi si otterranno numeri più probabili di altri, come il 7 che è il numero che si può ottenere con maggiori combinazioni possibili. Il gioco dei dadi è uno dei tentativi fatti, ancora prima della nascita del computer, per provare a comporre musica in maniera alternativa rispetto alla creatività umana.

	1	2	3	4	5	6	7	8
2	96	22	141	41	105	122	11	30
3	32	6	128	63	146	46	134	81
4	69	95	158	13	153	55	110	24
5	40	17	113	85	161	2	159	100
6	148	74	163	45	80	97	36	107
7	104	157	27	167	154	68	118	91
8	152	60	171	53	99	133	21	127
9	119	84	114	50	140	86	169	94
10	98	142	42	156	75	129	62	123
11	3	87	165	61	135	47	147	33
12	54	130	10	103	28	37	106	5

	1	2	3	4	5	6	7	8
2	70	121	26	9	112	49	109	14
3	117	39	126	56	174	18	116	83
4	66	139	15	132	73	58	145	79
5	90	176	7	34	67	160	52	170
6	25	143	64	125	76	136	1	93
7	138	71	150	29	101	162	23	151
8	16	155	57	175	43	168	89	172
9	120	88	48	166	51	115	72	111
10	65	77	19	82	137	38	149	8
11	10	4	31	164	144	59	173	78
12	35	20	108	92	12	124	44	131

Fig. 1. Tabelle del gioco dei dadi di Mozart. Il minuetto è diviso in due parti da 8.

Un tale sistema aleatorio, come si può intuire, è tutt'altro che creativo; questo infatti, anche se realizza un nuovo e inedito brano musicale, non dipende da proprietà musicali ma è strettamente legato a un "gioco" di calcolo combinatorio puramente matematico. Per questo motivo bisogna chiarire come distinguere un sistema creativo da uno non creativo che comunque genera novità e può essere utile a scopi artistici. Da qui la necessità di valutare non solo il prodotto, la creazione compiuta, ma anche il processo con cui essa è stata realizzata.

## 2.1 Evoluzione dell'intelligenza artificiale musicale creativa

Nel 1955 John McCarthy, informatico e scienziato cognitivo, conia il termine "intelligenza artificiale". Harun Zulić, citando Barr e Feigenbaum, definisce l'Intelligenza Artificiale come «quella branca dell'informatica che si occupa della progettazione di sistemi informatici intelligenti che presentano le caratteristiche che associamo all'intelligenza nel comportamento umano: comprensione del linguaggio, apprendimento, risoluzione di problemi e così via»<sup>11</sup>. È fondamentale precisare che l'intelligenza artificiale è un campo molto

<sup>11</sup> H. Zulić, *How AI can Change/Improve/Influence Music Composition, Performance and Education: Three Case Studies*, in «INSAM Journal of Contemporary Music, Art and Technology», fasc. 1, 2019, pp. 100-114:101.

vasto e include anche approcci e tecniche che non coinvolgono alcun apprendimento; quest'ultimo infatti è trattato nell'ambito del *machine learning*, che è solo una delle branche dell'intelligenza artificiale e non ne rappresenta la totalità. Il primo brano generato da un computer risale al 1957; si tratta di solo audio senza partitura ed è una melodia di 17 secondi dal titolo *The Silver scale*. Il titolo è stato attribuito da Newman Guttman, musicista, informatico e "padre" del progetto che ha condotto a questa prima composizione; la melodia è composta da Music I, un software per la sintesi sonora. Nello stesso anno viene composta da computer anche la partitura della *Illiac suite*, un brano per quartetto d'archi. Queste prime due composizioni non hanno ancora a che fare con l'apprendimento di un computer in quanto sono il frutto di modelli stocastici<sup>12</sup> (processi Markoviani, cfr. par. 2.3). Nel 1960, cinque anni dopo la definizione di McCarthy, inizia a svilupparsi l'idea di una possibile associazione di intelligenza artificiale e musica con la pubblicazione *An algorithmic description of process of music composition* di R. Kh. Zaripov<sup>13</sup>. Una delle prime concrete applicazioni di questa nuova branca della ricerca è un programma, sviluppato da Raymond Kurzweil nel 1963, in grado di riconoscere pattern in composizioni musicali già esistenti e utilizzarli per comporre nuova musica simile a quella dei compositori da cui il programma apprende. Il legame tra musica e computazione non rimane però confinato solo alla comunione tra musica e intelligenza artificiale, ma è esteso ad un generico sviluppo di programmi per supportare la creatività con modelli per la sintesi del suono e con software per assistere il compositore nel processo creativo automatizzando alcune parti del processo compositivo.

Intorno agli anni 80 iniziano ad essere sviluppati software CAD (computer aided design), cioè software utilizzati in varie discipline come architettura o il disegno di circuiti elettronici complessi, in grado di semplificare la fase di progettazione. In parallelo si inizia a pensare anche a un software CAC (computer aided composition) cioè un «sistema che dovrebbe fungere da strumento per facilitare la 'progettazione' di grandi strutture musicali»<sup>14</sup>. In questo secondo caso l'obiettivo del software è quello di automatizzare processi in cui il compositore può definire le regole sintattiche del linguaggio musicale che vuole utilizzare; a questo punto il computer applica queste regole per generare strutture musicali semplici. Grazie alla possibilità di effettuare più volte il processo si possono creare composizioni via via più complesse o articolate; il sistema infatti associa le strutture musicali semplici in sovrastrutture così come sono definite da S. R. Holtzman. I software di tipo CAC vengono sviluppati nel tempo seguendo due direzioni diverse. La prima segue una via sperimentale e di ricerca con compositori come Iannis Xenakis: in questo caso programmi di tipo CAC si sono

<sup>12</sup> J.-P. Briot *et al.*, *Deep Learning Techniques for Music Generation*, Springer International Publishing, Cham 2020, p. 2.

<sup>13</sup> Zulić, *How AI can Change/Improve/Influence Music Composition*, p. 101.

<sup>14</sup> Holtzman, *Generative grammars and the computer-aided composition of music*, University of Edinburgh, 1980, p. 5.

evoluti acquisendo molta autonomia nel senso in cui il compositore detta regole di tipo matematico stocastiche in base alle quali genera una successione di note o un'armonia. La seconda segue una via commerciale: in questo caso i software CAC si sono evoluti a software di produzione musicale per cui la parte creativa rimane in mano al compositore e il software fa da assistente e strumento atto a tenere in ordine la composizione. Software commerciali di tipo CAC sono ad esempio i DAW (Digital Audio Workstation), cioè software progettati per la composizione, il montaggio e la riproduzione di audio digitale. Queste interfacce permettono di raccogliere in un unico sistema tutti gli strumenti necessari per comporre musica con l'ausilio di un computer, da un editor della partitura all'esecuzione della stessa usando degli strumenti virtuali (plugin VST). Questi software, in linea con l'idea iniziale dei CAC, permettono di programmare il computer perché questo generi ad esempio delle progressioni, piuttosto che dei giri di accordi, secondo regole prestabilite dal compositore.

Alcuni software CAC si sono evoluti fino a sostituire quasi del tutto la figura del compositore permettendone l'utilizzo anche ai non addetti ai lavori e diventando il software stesso il compositore.

Un esempio eclatante in quanto primo 'compositore virtuale' le cui creazioni vengono registrate presso una società per i diritti d'autore, la SACEM (Société des Auteurs, Compositeurs et Editeurs de Musique), è AIVA (acronimo di Artificial Intelligence Virtual Artist), nata nel 2016. AIVA, così come definita da AIVA Technologies SARL, azienda che sviluppa AIVA, è «un'intelligenza artificiale che compone colonne sonore emozionanti per film, giochi, trailer e altri contenuti di intrattenimento»<sup>15</sup>. AIVA utilizza modelli di Deep Neural Network (cfr. par. 2.3) per modellizzare i concetti fondamentali della teoria musicale e l'arte del comporre. Come afferma H. Zulić, citando in maniera diretta un comunicato stampa di AIVA Technologies, AIVA serve per «rappresentare un modello matematico di cosa la musica sia. Questo modello viene poi utilizzato da AIVA per scrivere musica assolutamente unica»<sup>16</sup>. Il modello, costruito su un dataset di 30.000 composizioni in formato MIDI (cfr. par. 2.2), è in grado di riconoscere pattern nel modo di comporre nei diversi stili, per poi prevedere cosa deve venire dopo nella composizione rispetto a ciò che è stato precedentemente scritto. Inoltre AIVA possiede un sistema anti plagio che le permette di verificare che i passaggi delle sue creazioni non siano, per un errore di apprendimento del modello o per overfitting<sup>17</sup>, troppo simili o uguali ai passaggi delle composizioni su cui il modello è stato addestrato. Bisogna porsi però come quesito quale sia il vero scopo di AIVA e se un modello così accurato possa davvero sostituire l'estro creativo umano. Si può provare a trovare la risposta proprio in uno dei

<sup>15</sup> AIVA Press release, 2017 <https://www.luxinnovation.lu/wp-content/uploads/sites/3/2018/01/fundraising-en-aiva.pdf> (ultimo accesso: 01/09/23).

<sup>16</sup> Zulić, *How AI can Change/Improve/Influence Music Composition*, p. 104.

<sup>17</sup> Fenomeno che si verifica quando il modello si adatta troppo al set di dati usati per l'addestramento e non fornisce più previsioni accurate per i dati che non appartengono al dataset di addestramento ma che gli vengono forniti successivamente.

comunicati stampa di AIVA Technologies in cui l'azienda mette in chiaro quali siano i suoi obiettivi, cioè sviluppare un sistema in grado di comporre musica personalizzata e per lunghe ore di intrattenimento:

Gli sceneggiatori, i produttori cinematografici e l'intera industria dell'intrattenimento si affidano alla musica per trasformare momenti in magia. Ma i contenuti interattivi come i videogiochi hanno centinaia di ore di gioco e in genere non più di due ore di musica, in quanto i compositori umani sono limitati da vincoli fisici. Invece di riproporre sempre gli stessi brani, l'IA offre un'alternativa per aumentare le abilità creative umane in modo che i giochi possano finalmente avere centinaia di ore di musica originale.<sup>18</sup>

AIVA è funzionale quindi a comporre colonne sonore secondarie alla principale composta da un compositore umano, è funzionale alla produzione di musica originale per molte ore di gioco ed è funzionale a comporre musica personalizzata in funzione ascoltatore; nonostante questo non può sostituire del tutto la figura del compositore umano.

## 2.2 Forme di rappresentazione musicale digitale

La musica in forma digitale si può rappresentare in vari modi. Dalla scelta del tipo di rappresentazione dipendono le caratteristiche di una composizione che il computer sarà in grado di modellizzare.

La maniera più semplice di rappresentare una traccia audio è la forma d'onda del suono, il segnale audio puro. Questa rappresentazione non descrive propriamente la composizione, quindi una successione di note, armonie, dinamica, ma descrive l'ampiezza del segnale audio nel tempo e quindi contiene tutte le informazioni possibili, i dati grezzi, che si possono raccogliere in una registrazione audio. Se è vero che sistemi come il *Deep Learning* sono molto efficaci nell'elaborazione di dati grezzi non strutturati è anche vero che in questo caso, quando si parla di creatività e computazione, l'essenza della musica sta nel processo compositivo, che è rappresentato da simboli che vengono strutturati e ordinati su una partitura, cosa che la forma d'onda sonora non può rappresentare<sup>19</sup>. Parlando di simbologia musicale è utile ripercorrere quali sono i simboli atti ad una rappresentazione musicale comprensibile da un computer<sup>20</sup>:

– *altezza del suono*: può essere descritta in Hertz, quindi come frequenza oppure con una notazione che combini il nome della nota con un numero che identifica l'ottava a cui la nota appartiene (ad esempio A4 rappresenta il la a 440Hz)

– *durata*: può essere descritta in termini assoluti in millisecondi oppure con valore relativo, quello della notazione musicale classica (es: un quarto)

<sup>18</sup> AIVA Press release, cit.

<sup>19</sup> Briot et al., *Deep Learning Techniques for Music Generation*, p. 21.

<sup>20</sup> Ivi, p. 25; R. B. Dannenberg, *Music Representation Issues, Techniques, and Systems*, in «Computer Music Journal», fasc. 17, 3 1993, pp. 20-30:22.



- *dinamica*: può essere descritta in termini assoluti e quantitativi in decibel (dB) oppure con valore qualitativo, quello della notazione musicale classica (es: pp)
- *pausa*: può essere descritta in termini assoluti in millisecondi oppure con valore relativo, quello della notazione musicale classica (es: pausa da un quarto)

Un formato utile a rappresentare una composizione musicale che sia comprensibile da un computer è il formato MIDI (Musical Instrument Digital Interface). questo formato, mediante l'utilizzo di un protocollo di comunicazione<sup>21</sup>, riesce a governare l'interazione in tempo reale tra i diversi strumenti musicali elettronici e questi con un computer ad essi connesso. Il MIDI codifica in maniera chiara quelle informazioni scritte sopra, cioè altezza del suono, durata, dinamica e pausa, utili a rappresentare una composizione.

Le istruzioni di un file MIDI, chiamate anche 'eventi', vengono codificate in righe da 16 valori esadecimali a cui si aggiunge un'ultima riga di 17 valori. Un evento midi può essere ad esempio il suonare una nota. In questo caso l'evento *Note On* è rappresentato nella seguente maniera: '9n hh vv' dove '9' sta per lo status byte, cioè l'evento Note On; 'n' è un numero che identifica il canale, cioè o lo strumento o la traccia; 'hh' è un numero che rappresenta la nota che suona; 'vv' è la velocità di tocco, cioè la forza con cui la nota viene suonata e quindi il volume della nota.

Nella stessa maniera si possono rappresentare il rilascio della nota, pause, metrica, dinamica, si possono sovrapporre note per rappresentare accordi e molto altro<sup>22</sup>. Questo permette una visualizzazione della composizione utile a fare sì che il computer identifichi delle caratteristiche e possa trovare dei pattern in un set di addestramento (un insieme di composizione su cui il computer viene allenato a comporre).

Oltre all'utilizzo del formato MIDI è possibile rappresentare una melodia in molti modi diversi. Uno di questi, ad esempio, è la rappresentazione testuale che può essere compresa da un software appositamente sviluppato; in questo caso la notazione utilizzata è quella anglosassone, dove C4 corrisponde al do centrale, sarà poi il software in grado di interpretare il testo in questione<sup>23</sup>.

## 2.3 Metodi di composizione artificiale

I metodi per comporre musica in maniera artificiale sono molteplici e spesso vengono usati insieme per arrivare al prodotto finale. Alcuni aspirano ad essere definibili come creativi, quindi ad essere associati ad un processo creativo che tenta di emulare quello umano, altri sono soluzioni complesse

<sup>21</sup> Un protocollo di comunicazione in informatica è un insieme di regole che permettono il corretto scambio di informazioni tra due dispositivi. In questo caso il protocollo MIDI permette lo scambio di dati di tipo MIDI. Questi possono essere raccolti in file dall'omonimo nome ed essere serializzati.

<sup>22</sup> C. Giustozzi, *Il protocollo MIDI*, in «MCmicrocomputer», fasc. 127, 1993, pp. 194-196.

<sup>23</sup> Briot et al., *Deep Learning Techniques for Music Generation*, pp. 32-33.

per cercare di modellizzare il processo di composizione musicale, senza forzatamente aspirare ad una componente creativa di tipo umano. Di seguito vengono presentati alcuni tra i metodi che F. Carnovalini e A. Rodà identificano tra i più popolari<sup>24</sup>:

– *Processi Markoviani*: i processi Markoviani sono processi stocastici che descrivono la probabilità di una sequenza di stati. Quando si parla di catene Markoviane si intende per stato un oggetto variabile che può assumere diverse identità. Un esempio può essere identificare uno stato con una nota, così come identificarlo nella durata della nota o in una qualsiasi struttura musicale. In una catena Markoviana uno stato dipende solo dallo stato precedente e non da come si è arrivati a quest'ultimo. Il modo più semplice di implementare un modello Markoviano è quello di identificare ogni stato con una nota del brano. In questo modo il computer andrà a calcolare quale è la probabilità che a una nota segua un'altra nota. Un modo più complesso di utilizzare le catene Markoviane è quello di affiancare catene per la melodia con catene per la durata delle note. Esistono diversi approcci da utilizzare in funzione della specifica situazione. Da un punto di vista creativo le catene Markoviane rischiano di riutilizzare molto materiale dal set di addestramento in maniera non creativa, anche plagiando le composizioni da cui ha appreso. Un esempio dell'utilizzo di questo modello è la *Illiac suite* (cfr. par. 2.1).

– *Grammatica formale e generativa*: la grammatica formale è un modello matematico che identifica in una lingua un insieme infinito di sequenze finite di simboli, quindi di frasi, appartenenti ad un alfabeto. Rifacendoci alla categorizzazione della creatività di Boden (cfr. par. 1.2), questo insieme può essere definito lo spazio concettuale di una lingua. La grammatica generativa è invece un insieme di regole di riscrittura secondo cui le sequenze finite di simboli vengono combinate per creare altre frasi<sup>25</sup>. Il concetto di grammatica generativa viene sviluppato da Noam Chomsky come strumento per l'analisi del linguaggio naturale. F. Lerdahl e R. Jackendoff applicheranno la stessa idea alla musica trasformando i concetti di grammatica formale e grammatica generativa in strumenti per l'analisi musicale<sup>26</sup>. Per fare un esempio dell'applicazione al campo musicale si prenda in considerazione il serialismo sviluppato da Schoenberg: in questo caso la grammatica formale si riferisce alle dodici note sviluppate in successione secondo una 'serie', che corrisponde alla grammatica generativa, cioè regole secondo cui le dodici note sono organizzate lungo tutto il brano.

Utilizzare una grammatica formale e una generativa in campo musicale significa trovare regole grammaticali, relative a una grammatica musicale, secondo

<sup>24</sup> Carnovalini, Rodà, *Computational Creativity and Music Generation Systems: An Introduction to the State of the Art*, p. 8.

<sup>25</sup> S. R. Holtzman, *Using Generative Grammars for Music Composition*, in «Computer Music Journal», fasc. 5, 1 1981, pp. 51-64:52.

<sup>26</sup> Carnovalini, Rodà, *Computational Creativity and Music Generation Systems*, p. 9.

cui i simboli musicali in una composizione vengono messi in successione per dare forma alla composizione stessa. In questo caso il simbolo musicale non ha lo stesso significato che ha nella notazione musicale classica ma si può definire come qualsiasi elemento possa essere parametrizzato in una composizione, quindi una nota ma anche un accordo o una progressione o un intervallo, ciò che fa più comodo parametrizzare in funzione del caso specifico. Uno dei passaggi fondamentali nell'utilizzo di questa tecnica è definire un adeguato alfabeto di simboli che possano essere organizzati in sequenze per un numero di sequenze infinite e secondo regole sempre applicabili e universali.

– *Rule/Constraint based system*: anche in questo caso, come accade per il metodo precedente, si ha una serie di regole applicate a sequenze di variabili (quelle che nella grammatica formale vengono chiamate simboli). In questo caso però non sono le regole ad essere costruite in funzione di un alfabeto, ma sono le sequenze di variabili ad essere cercate da un dominio (l'insieme dei simboli musicali) perché possano essere messe in ordine con delle regole e dei vincoli precedentemente definiti<sup>27</sup>.

Tornando alla classificazione dei tipi di creatività di Boden l'integrazione in un software di queste regole può essere vista in due modi<sup>28</sup>: il primo è considerare le regole come trasformative dello spazio concettuale introdotto da Boden, che in questo caso corrisponde all'insieme dei simboli musicali che formano il dominio entro cui queste regole scelte vengono applicate; il secondo è considerarle come indicative per l'esplorazione dello spazio concettuale (cfr. par. 1.2).

– *Reti neurali/Deep Learning*: tra quelle presentate in questo articolo il *Deep Learning* è il primo metodo che comporta un'effettiva fase di apprendimento da parte di un computer. Il *Deep Learning* è una sottocategoria del *Machine Learning*, cioè quell'insieme di sistemi che non vengono propriamente programmati ma vengono addestrati. Infatti se nelle altre tecniche le regole per comporre vengono previste e scritte già in fase di scrittura del software, in un sistema di *Deep Learning* le regole e i vincoli vengono definiti in maniera autonoma dal computer. Un modello di *Deep Learning* fa uso di reti neurali artificiali. Queste reti sono modelli matematici composte da nodi (neuroni) ciascuno dei quali può ricevere più input, cioè informazioni da altri nodi o dall'esterno della rete, e genera un solo output, che è elaborazione degli input ricevuti, che può essere il risultato della rete o diventare input di altri nodi. Una rete neurale è composta da più layer, livelli, di neuroni come in Fig. 2.

<sup>27</sup> Ö. Sandred, PWWC, *a Constraint-Solving System for Generating Music Scores*, in «Computer Music Journal», fasc. 34, 2 2010, pp. 8-24:9.

<sup>28</sup> Carnovalini, Rodà, *Computational Creativity and Music Generation Systems*, p. 10.

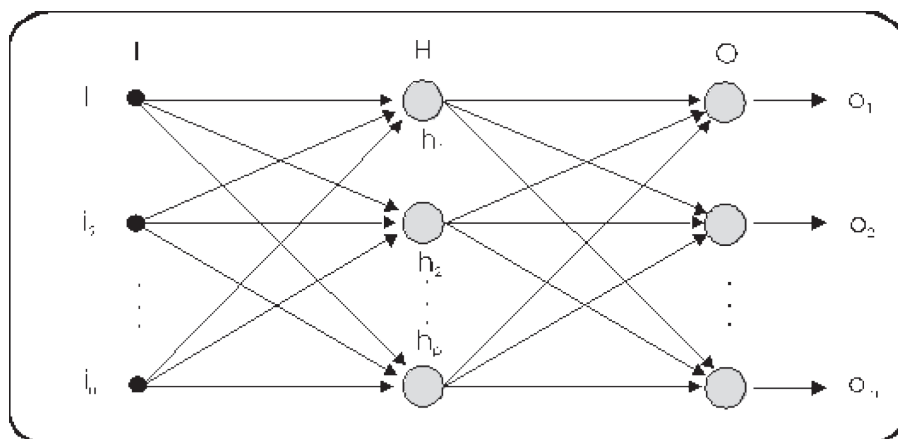


Fig. 2. Rete neurale di un modello di *Deep Learning*.  
Immagine di dominio pubblico via Wikimedia Commons.

In Fig. 2 I e O sono rispettivamente i layer di input e output, quindi il layer da cui arrivano i dati alla rete e il layer con la risposta della rete. H è chiamato hidden layer, cioè un layer interno alla rete i cui input arrivano dal layer precedente e i cui output vanno al layer successivo. il ruolo del layer è quello di raffinare un dato di input in modo che l'output contenga solo le caratteristiche di quel dato che la rete ritiene utili. In un modello di *Deep Learning* possono esserci anche centinaia di layer. Le strutture dei layer vengono modificate in autonomia dal modello secondo regole matematiche non lineari, in modo che la rete possa gestire anche dati solitamente non operabili in matematica, confrontando quella che è la previsione della rete con quella che dovrebbe essere la previsione corretta contenuta in un dataset di addestramento e modificandosi di conseguenza. Questo è il motivo per cui non è possibile conoscere il processo secondo cui una rete neurale quando viene utilizzata compone un brano. Nella programmazione classica il programmatore scrive un insieme di regole e inserisce nel software dei dati da elaborare in base a queste regole ottenendo delle risposte. Con il *machine learning* il programmatore inserisce i dati e le risposte attese sulla base di questi dati ed è il computer ad individuare le regole, regole che risultano incomprensibili all'occhio umano. «Un sistema di *machine learning* viene addestrato non programmato»<sup>29</sup>.

Un'evoluzione del normale approccio al *Deep Learning* è quella delle *Generative Adversarial Networks* (GANs). Questo approccio consiste nell'addestrare contemporaneamente due reti neurali: una avrà il compito di generare qualcosa di nuovo, l'altra avrà il compito di valutare se un dato è reale o generato dalla

<sup>29</sup> F. Chollet, *Deep learning con Python: imparare a implementare algoritmi di apprendimento profondo*, Apogeo, Milano 2020, p. 5.

prima rete. In questa maniera le due reti si possono migliorare a vicenda fino a quando i risultati della prima saranno quasi indistinguibili da un dato reale<sup>30</sup>.

Un modello di *Deep Learning*, e più in generale uno di *Machine Learning*, sono più funzionali e più vicini a quello che è lo scopo presentato all'inizio di questo testo, quindi quello di realizzare un computer creativo in grado di emulare la creatività umana, rispetto a tecniche come la grammatica generativa o i *Rule based system* perché un «sistema generativo basato sull'apprendimento automatico può essere agnostico, in quanto impara un modello da un corpus musicale arbitrario. Di conseguenza lo stesso sistema può essere utilizzato per diversi generi musicali»<sup>31</sup>. Un esempio concreto di quella che è la potenzialità di un modello di *Deep Learning* è AIVA, di cui si è parlato nel paragrafo 2.1.

– *Algoritmi genetici*: ispirato all'evoluzione biologica un algoritmo genetico parte da una *popolazione* casuale di soluzioni al compito dato e cerca di combinare e selezionare le soluzioni migliori per ottenerne di nuove avvicinandosi sempre di più alla soluzione ottimale<sup>32</sup>. Ad ogni ciclo evolutivo l'algoritmo apporta delle piccole modifiche agli *individui* della popolazione per migliorarli e tiene memoria, per i cicli evolutivi successivi, della reazione della *popolazione* alle modifiche. Questo comporta una forte capacità di adattamento al compito, «capacità di adattamento possibile grazie all'uso di strutture di controllo, feedback, dove un cambiamento futuro tiene conto di come cambiamenti simili abbiano influenzato sistemi simili in passato»<sup>33</sup>. Un esempio di applicazione in campo musicale di un algoritmo genetico è GenJam, software sviluppato da Al Biles, che permette al computer di improvvisare in tempo reale musica jazz interagendo con un musicista umano che improvvisa a sua volta e ascoltando il musicista in modo da potere rispondere. Le popolazioni a cui GenJam fa riferimento sono due: una relativa alle frasi musicali e una al tempo<sup>34</sup>. Un limite degli algoritmi generici è però che, ad oggi, essi necessitano di un operatore umano che imponga dei parametri iniziali secondo cui la popolazione viene valutata.

## 2.4 Possibili sviluppi

Sono molte le direzioni che la ricerca sta prendendo nel campo dell'intelligenza artificiale creativa. Una che merita attenzione sia in quanto nuova prospettiva che come alternativa alla linearità dei sistemi computazionali tradizionali è l'informatica quantistica. Un computer quantistico sfrutta la meccanica quantistica, quindi la descrizione di sistemi fisici a livello subatomico, per aumentare note-

<sup>30</sup> Carnovalini, Rodà, *Computational Creativity and Music Generation Systems*, p. 10.

<sup>31</sup> Briot *et al.*, *Deep Learning Techniques for Music Generation*, p. 5.

<sup>32</sup> Carnovalini, Rodà, *Computational Creativity and Music Generation Systems*, p. 11.

<sup>33</sup> C.G. Johnson, J.J.R. Cardalda, *Genetic Algorithms in Visual Art and Music*, in «Leonardo», fasc. 35, 2 2002, pp. 175-184:177.

<sup>34</sup> J.A. Biles, *GenJam: A Genetic Algorithm for Generating Jazz Solos*, in «ICMC», fasc. 94, 1994, pp. 131-137:132.

volmente la potenza di calcolo del processore di un computer. Oltre alla velocità i computer quantistici hanno un'altra importante caratteristica: se nei computer tradizionali un bit può assumere come valore 0 o 1 nei computer quantistici un qubit può essere contemporaneamente sia 0 che 1. Questo permette di uscire da una computazione lineare aprendo come possibilità quella di riprodurre la molteplicità di stati che la mente può assumere contemporaneamente. Questo prima era impossibile in quanto i computer tradizionali sono limitati dalla sola possibilità di lavorare solo con 0 e 1 in successione lineare e non in parallelo, come accade appunto con i computer quantistici.

## Conclusioni

Negli ultimi decenni abbiamo assistito a una rapida evoluzione nella creatività computazionale musicale grazie all'uso di algoritmi di apprendimento automatico, reti neurali artificiali e tecnologie avanzate di elaborazione del suono. Questi sviluppi hanno consentito alle macchine di generare musica in una varietà di stili e generi in modi che sfidano le aspettative tradizionali. Nonostante i progressi, la definizione stessa di creatività musicale è al centro del dibattito odierno. La creatività è spesso vista come un processo umano intrinsecamente legato alle emozioni, all'originalità e all'espressione personale. Tuttavia, le macchine non provano emozioni come gli esseri umani, il che solleva domande sulla vera natura della creatività nelle opere generate da un computer. Per decretare se una macchina sia creativa o meno bisogna trovare un accordo comune sul significato di creatività. La comunità scientifica propone diverse teorie e studi per comprendere cosa la creatività sia ma ancora non si hanno conoscenze tali sul funzionamento della mente umana per definire in maniera ontologica la sua essenza e da cosa essa dipenda. Da questo deriva la molteplicità di metodi per provare a emulare la creatività in maniera artificiale, ma, poiché ancora non si sa esattamente cosa la creatività sia per l'uomo, quello di sviluppare sistemi artificialmente intelligenti rimane un tentativo di cui non è possibile verificare effettivamente la riuscita. Al di là dei costrutti e riconoscendo come cardine la multifattorialità del concetto di creatività, essa stessa potrebbe essere, in buona sostanza, considerata come la capacità di sorprendere, emozionare e ispirare gli ascoltatori. Le definizioni possono però variare ampiamente a seconda del contesto culturale e personale. L'importanza della creatività risiede nella sua capacità di trasmettere un significato profondo e di connettersi con le persone a livello emotivo, qualcosa che può essere valutato sia negli esperimenti musicale umani che nelle composizioni generate da un computer. Abbiamo visto come la ricerca si è focalizzata principalmente sulla dimensione cognitiva del processo creativo (cfr. par. 1.1) per la versatilità con cui essa si presta ad essere modellizzata ed imitata da un computer. Per questo un aspetto interessante della creatività computazionale in musica è la possibilità di una collaborazione tra umani e macchine. Questo si traduce in una sintesi creativa che sfrutta le forze complementari delle due parti. Gli umani contribuiscono con la visione, mentre

le macchine portano conoscenza, analisi dati avanzata e capacità di generazione veloce. Questa collaborazione uomo-macchina potrebbe portare a risultati unici e sorprendenti che altrimenti sarebbero difficili da ottenere. Rimane invece ancora da esplorare ampiamente come e se la dimensione emotiva della creatività possa essere riprodotta da un computer, e se questo sia davvero necessario per potere generare artificialmente un prodotto definibile come creativo. Una domanda da porsi è se la componente emotiva di un brano possa essere ridotta solo a percezione dell'ascoltatore. Ci si può emozionare anche davanti a una composizione scritta da una macchina? L'emotività dipende solo dal compositore o può essere intrinseca alla tecnica compositiva, e quindi può essere modellizzata dal computer? Si può ridurre la mente umana a un sistema molto complesso che presenta come proprietà emergente una dimensione emotiva ma che nelle sue singole parti è riducibile a modello matematico? Queste sono le domande a cui la comunità scientifica dovrà dare una risposta per arrivare alla teorizzazione completa di un sistema creativo.

In conclusione, lo stato attuale della creatività computazionale in musica offre numerose opportunità ma presenta anche sfide profonde. Queste sfide sono radicate nella stessa definizione di creatività e nella capacità di comunicare emozioni attraverso la musica generata da un computer. La collaborazione tra esseri umani e macchine emerge come un processo promettente per superare queste sfide e scoprire nuove frontiere creative in campo musicale. È in questa sintesi tra la creatività umana e quella dell'IA che il futuro della musica potrebbe essere ridefinito, offrendo emozionanti possibilità sia per gli artisti che per gli ascoltatori. Mentre la tecnologia continua a evolversi, l'esplorazione di queste intersezioni creative ha potenziale per ridefinire la nostra comprensione di ciò che è veramente possibile nel mondo della musica.