

**AfIA**Association française
pour l'Intelligence Artificielle

■ Monte Carlo Tree Search pour la politique monétaire : une approche par planification séquentielle

LAMSADE

Université Paris Dauphine-PSL

Université Paris Dauphine-PSL

Tristan CAZENAVEtristan.cazenave@lamsade.dauphine.fr**Mouhssine RIFAKI**mouhssine.rifaki@dauphine.eu

Introduction

Nous proposons d'explorer l'application de Monte Carlo Tree Search (MCTS) à la politique monétaire. Cette approche novatrice pourrait potentiellement améliorer les décisions de politique monétaire par rapport aux règles traditionnelles, particulièrement en situation de crise.

Les banques centrales utilisent principalement des règles de politique monétaire comme la règle de Taylor et ses variantes. Ces règles, basées sur des formules linéaires, peuvent présenter des limites face à des situations économiques complexes : contraintes de taux zéro, crises financières, chocs d'offre importants.

Notre approche propose d'utiliser des méthodes de planification séquentielle issues de l'intelligence artificielle. Au lieu de suivre une formule préétablie, nous transformons chaque décision de taux d'intérêt en problème de planification où l'algorithme simulerait des milliers de trajectoires économiques possibles pour évaluer différentes options de politique monétaire.

Objectifs de notre approche DSGE-MCTS :

**Améliorer les décisions de politique
monétaire en situation complexe**

**Explorer systématiquement
l'incertitude future**

**Fournir un outil complémentaire aux
règles traditionnelles**

Architecture DSGE-MCTS : fusion de deux approches

Notre proposition combine deux approches complémentaires :

1. Monte Carlo Tree Search (MCTS) : Un algorithme de planification qui a montré son efficacité dans des domaines complexes comme les jeux de stratégie [3, 1]. MCTS explore efficacement des espaces décisionnels vastes en présence d'incertitude.

2. Dynamic Stochastic General Equilibrium (DSGE) : Les modèles macroéconomiques standards utilisés pour l'analyse de politique monétaire. Ces modèles capturent les interactions entre agents économiques dans un cadre d'équilibre général.

Framework proposé. Le système fonctionnerait en quatre étapes :

Étape 1 - Représentation d'état : À chaque trimestre t , l'état $s_t = (x_t, z_t, E_t x_{t+1}, \Theta_t)$ encoderait les variables endogènes (inflation, production), les chocs exogènes (productivité, demande), les anticipations rationnelles et les statistiques du nœud.

Étape 2 - Espace d'action : Le taux d'intérêt nominal $a_t \in \mathcal{A}$ serait discrétisé en **41**



points espacés de 5 points de base au nœud racine (décision réelle). Dans l'arbre de recherche, pour économiser les ressources computationnelles, chaque nœud interne commencerait avec seulement 5 actions. L'élargissement progressif (progressive widening) [2] ajouterait dynamiquement de nouveaux points selon $N(s)^{1/3}$ à mesure que le nœud est visité, permettant d'explorer progressivement l'espace d'actions dans les branches prometteuses.

Étape 3 - Fonction de récompense : Le bien-être social instantané serait défini comme :

$$u(s_t, a_t) = \Delta C_t - \frac{\lambda_\pi}{2} \pi_t^2 - \frac{\lambda_y}{2} y_t^2$$

avec $(\lambda_\pi, \lambda_y) = (1.5, 0.1)$ suivant la calibration de Smets-Wouters [8].

Étape 4 - Dynamique de transition : Le modèle DSGE générerait $s_{t+1} = \mathcal{F}(s_t, a_t, \varepsilon_{t+1})$ via la solution OccBin [4] pour gérer les contraintes non-linéaires (taux zéro).

L'algorithme de planification proposé

À chaque décision trimestrielle, DSGE-MCTS exécuterait $M = 5000$ simulations Monte Carlo structurées par l'heuristique PUCT [6, 7] :

$$UCB(s, a) = \bar{Q}(s, a) + c_{PUCT} \times P(s, a) \times \frac{\sqrt{\sum_b N(s, b)}}{1 + N(s, a)}$$

où $\bar{Q}(s, a)$ est la valeur moyenne estimée, $P(s, a)$ le prior généré par un réseau de neurones, et $c_{PUCT} = 1.5$ la constante d'exploration.

Architecture envisagée du réseau de neurones pour le prior $P(s, a)$

ENTRÉES : L'état économique s_t

(vecteur de dimension d) contenant :

- Variables endogènes (inflation, output gap, consommation...)
- Chocs exogènes (productivité, demande...)
- Anticipations rationnelles des agents
- Statistiques des nœuds de l'arbre

ARCHITECTURE : Réseau feedforward à 3 couches

256 neurones $\rightarrow$ 128 neurones $\rightarrow$ 64 neurones $\rightarrow$ sortie

Activation ReLU entre les couches

SORTIES : Probabilités $P(s, a)$ pour chaque action a possible (distribution de probabilité sur les 41 taux d'intérêt au nœud racine)

Le réseau serait entraîné hors-ligne sur des états de simulation pour prédire les meilleures actions selon l'historique des simulations. L'architecture serait sélectionnée après tests comparatifs sur plusieurs configurations.

Caractéristique notable : prise en compte des risques futurs. Contrairement aux règles traditionnelles qui réagissent principalement aux conditions courantes, MCTS évaluerait systématiquement les conséquences futures des décisions. L'algorithme pourrait ainsi détecter l'augmentation du risque de récession et ajuster préventivement la politique monétaire.

Propriétés de convergence attendues. Les méthodes de Monte Carlo Tree Search bénéficient de propriétés de convergence bien établies dans la littérature. Dans le contexte DSGE-MCTS, nous anticipons que l'algorithme convergerait vers les décisions optimales avec un nombre suffisant de simulations, sous



Afia

Association française
pour l'Intelligence Artificielle

réserve de conditions techniques appropriées sur la structure du modèle économique sous-jacent.

Avantages potentiels de l'approche

Gestion des non-linéarités. MCTS pourrait prendre en compte explicitement les régimes futurs de taux zéro dans son arbre de recherche, contrairement aux approches linéaires-quadratiques traditionnelles.

Exploration pondérée par le risque. L'échantillonnage Monte Carlo permettrait une meilleure évaluation des scénarios de queue de distribution [5], particulièrement important en période de crise.

Adaptabilité. L'approche pourrait s'adapter automatiquement à différents régimes économiques sans nécessiter de changement de règle explicite.

Extensions et perspectives

Multi-agents et coordination internationale. Une extension intéressante consisterait à explorer le cas où plusieurs banques centrales utiliseraient DSGE-MCTS simultanément, permettant potentiellement une meilleure coordination des politiques monétaires.

Apprentissage des paramètres structurels. Le framework pourrait s'adapter à l'apprentissage en ligne des paramètres du modèle DSGE via maximum de vraisemblance sur données temps réel.

Interface pour analyse de scénarios. Un prototype pourrait permettre aux décideurs d'explorer l'arbre de décision et de tester différents scénarios économiques de manière interactive.

Défis et considérations pratiques

L'implémentation de cette approche présenterait plusieurs défis :

- **Ressources computationnelles** : L'approche nécessiterait des capacités de calcul importantes pour être utilisée en temps réel.
- **Calibration et validation** : La performance dépendrait fortement de la spécification et calibration du modèle DSGE sous-jacent.
- **Interprétabilité** : Il serait crucial de développer des outils permettant aux décideurs de comprendre les recommandations de l'algorithme.
- **Tests extensifs** : Une adoption pratique nécessiterait une phase de test approfondie en parallèle des méthodes existantes.

Message clé : Les méthodes de planification séquentielle pourraient offrir de nouvelles perspectives prometteuses pour l'analyse de politique monétaire, en complément des approches traditionnelles.

Références

- [1] Cameron B. Browne et al. A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in Games*, 4(1) :1–43, 2012.
- [2] Adrien Couëtoux, Jean-Baptiste Hoock, Nataliya Sokolovska, Olivier Teytaud, and Nicolas Bonnard. Continuous upper confidence trees. In *Learning and Intelligent Optimization : 5th International Conference, LION 5, Rome, Italy, January 17–21, 2011. Selected Papers 5*, pages 433–445. Springer, 2011.
- [3] Rémi Coulom. Efficient selectivity and backup operators in monte-carlo tree search. In *International Conference on Computers and Games*, pages 72–83, 2006.



AfIA

Association française
pour l'Intelligence Artificielle

- [4] Luca Guerrieri and Matteo Iacoviello. Occbin : A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics*, 70 :22–38, 2015.
- [5] Conor F. Hayes, Mathieu Reymond, Diederik M. Roijers, Enda Howley, and Patrick Mannion. Monte carlo tree search algorithms for risk-aware and multi-objective reinforcement learning. *Autonomous Agents and Multi-Agent Systems*, 37(2) :26, 2023.
- [6] David Silver et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529 :484–489, 2016.
- [7] David Silver et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv :1712.01815*, 2017.
- [8] Frank Smets and Raf Wouters. Shocks and frictions in us business cycles : A bayesian dsge approach. *American Economic Review*, 97(3) :586–606, 2007.

Contact pour collaboration : Les auteurs recherchent des collaborations avec les institutions académiques et les banques centrales intéressées par l'exploration de cette approche novatrice.