



Deep reinforcement learning

quand les robots apprennent à apprendre

Si actuellement les robots industriels ou les robots domestiques ne font que ce que leur programmation leur permet de faire, demain, ceux-ci vont réellement apprendre en fonction de ce qu'ils observeront. Une technique d'intelligence artificielle leur donne cette capacité à apprendre, c'est le *deep reinforcement learning*.

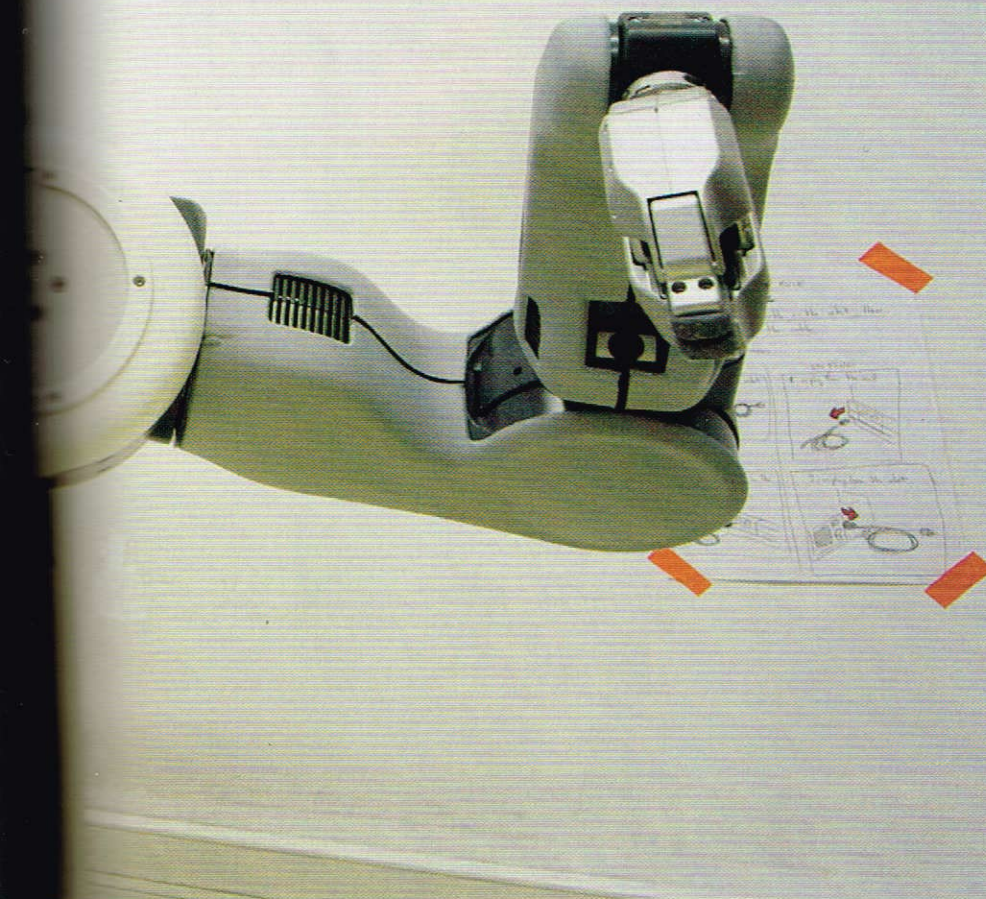
L'intelligence artificielle connaît ces derniers mois un renouveau sans précédent. Des algorithmes pourtant anciens ont retrouvé de l'intérêt du fait de la puissance des processeurs et surtout des cartes graphiques modernes, idéales pour porter des réseaux de neurones. La robotique bénéficie de cette vague notamment dans le domaine de l'analyse des images. « Le *deep learning* a fait ses preuves dans le domaine de l'analyse de l'image, et il commence à y

avoir des résultats de recherche extrêmement impressionnants dans ce que l'on appelle l'apprentissage par renforcement. C'est notamment le cas dans le domaine des jeux vidéo ou même des jeux traditionnels. Tout le monde se souvient de la réussite de l'équipe Google DeepMind avec AlphaGo et ses successeurs dans l'apprentissage du jeu de go » explique Stéphane Doncieux, responsable de l'équipe AMAC de l'ISIR (Institut des systèmes intelligents et de robotique) et coordinateur scientifique du projet européen DREAM.

UNE I.A. QUI APPREND PAR ESSAI-ERREUR

Désormais, chercheurs et industriels veulent aller plus loin, c'est-à-dire permettre aux algorithmes aux robots d'apprendre par eux-mêmes. Les experts en intelligence artificielle appellent cette technique le *deep reinforcement learning* ou apprentissage par renforcement. L'idée est de réaliser à un réseau de neurones une multitude d'essais pour résoudre un problème donné. Un système de récompense lui permet de savoir

Le projet européen *DREAM* (*Deferred Restructuring of Experience in Autonomous Machines*) cherche à doter les robots de capacités d'apprentissage avancées.



Nous sommes à une étape technologique extrêmement stimulante !

Stéphane Doncieux, responsable de l'équipe AMAC de l'ISIR (Institut des systèmes intelligents et de robotique) et coordinateur scientifique du projet européen *DREAM*

« L'objectif à terme de nos travaux est de pouvoir placer un robot dans une pièce totalement inconnue et que celui-ci soit capable, au travers de ses interactions de découvrir les objets, comprendre comment les manipuler sans assistance extérieure de façon à ce qu'on puisse donner une tâche à réaliser au robot et que celui-ci puisse la réaliser sans programmation. »

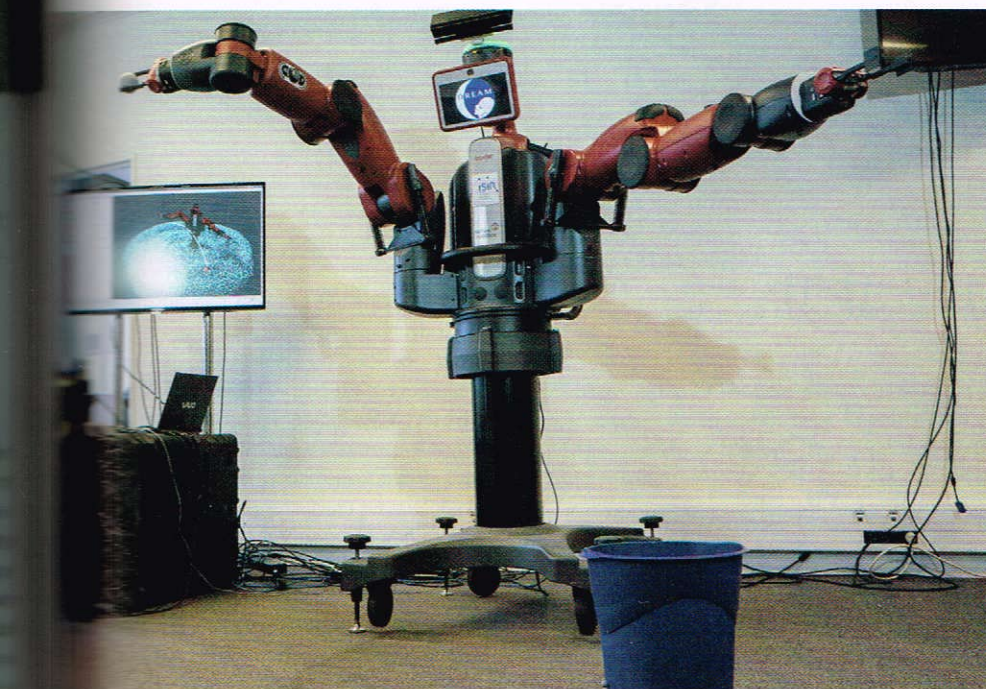
« Nous sommes aujourd'hui à une étape technologique extrêmement stimulante car de nombreux outils algorithmiques apparaissent, notamment ceux liés au deep learning entre autres. Ce dont nous avons le plus besoin aujourd'hui, c'est de tester différentes approches, pour sélectionner les plus appropriées. »

est plus ou moins pertinent, à la manière de la souris de laboratoire qui reçoit un bout de fromage lorsqu'elle trouve la sortie d'un labyrinthe. Cédric Buche, chercheur au Lab-STICC du CNRS et responsable du programme transverse intelligence artificielle du laboratoire informatique du Grand Ouest explique : « On pratiquait l'apprentissage par renforcement depuis très longtemps, mais le couplage du renforcement au deep learning devient de plus en plus fréquent comme on l'a vu avec Google. Cette approche est très puissante pour résoudre un certain nombre de problèmes. »

La technique a déjà démontré sa pertinence dans un domaine plutôt surprenant, celui des jeux. L'algorithme peut apprendre à jouer en jouant des centaines de milliers de parties. Dès 2013, une start-up anglaise nommée DeepMind dévoilait Atari Deep Q-Learner, un algorithme d'apprentissage par renforcement, capable de jouer à plusieurs jeux Atari dont les célèbres Breakout et Pong. Peu après cette démonstration, Google achetait DeepMind pour 400 millions de dollars. L'aboutissement de cette approche était la mise au point d'AlphaGo, l'intelligence artificielle qui allait vaincre le champion du monde du jeu de go en 2016. De Mario Bros. à Grand Theft Auto V en passant par Asteroids, de très nombreux jeux vidéo ont été vaincus par des I.A. et, récemment, il n'a fallu que quatre heures d'apprentissage pour que AlphaGo, rebaptisé AlphaZero AI, batte à plate couture le meilleur logiciel d'échecs du moment, Stockfish 8.

DES DIZAINES DE MILLIERS D'ESSAIS POUR RÉALISER UN APPRENTISSAGE

Appliquer cette technique permet à un robot d'accomplir un geste, mais ce passage de l'environnement totalement virtuel du jeu à la réalité n'est pas aussi simple qu'il n'y paraît. En effet, il faut faire des

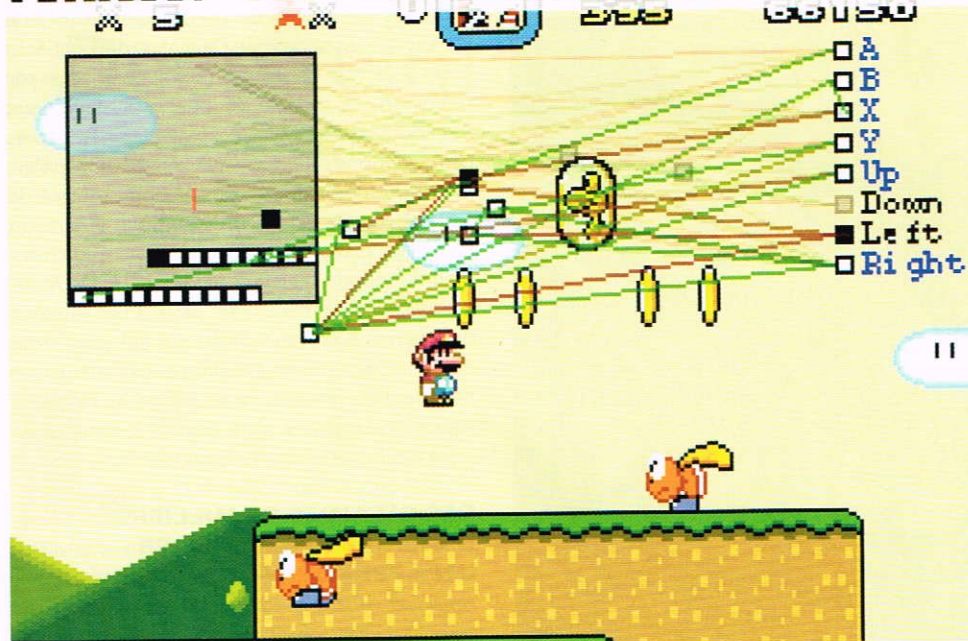




Le **deep learning** est une technique d'I.A. aujourd'hui parfaitement maîtrisée dans le domaine de l'analyse d'image.

Gen 34 species 14 genome 14 (37%)

Fitness: 381 Max Fitness: 4322



Le jeu vidéo est l'environnement idéal pour faire apprendre à un réseau de neurones à jouer, ici l'I.A. du chercheur Deepak Pathak de l'université de Berkeley s'attaque à Mario Bros.

milliers d'essais afin d'obtenir un apprentissage par renforcement correct. Ainsi, pour entraîner son robot gardien de but pour la RoboCup Football, l'équipe de chercheurs a dû tirer dix mille fois au but en simulation avant d'enfin tester son robot sur le terrain : « C'est qu'après cet apprentissage en simulation que le modèle appris a pu être transféré sur le robot physique pour tester et l'adapter au réel. »

En 2016, Google a mis en place une batterie de quatorze robots afin d'apprendre à un réseau de neurones à saisir des objets divers. Les trois mille heures d'utilisation accumulées par ces robots ont permis de réaliser huit cent mille essais. Beaucoup de laboratoires ont les moyens de mettre en place de telles installations. Mihai Polceanu, chercheur post-doc au LAB-STICC souligne : « L'apprentissage n'est pas faisable avec de vrais robots donc on préfère utiliser un simulateur avec lequel on simule le comportement du robot avec tous ses capteurs et axes de mouvement. On laisse tourner le simulateur sur un cluster de cartes graphiques pour obtenir des comportements qui semblent intéressants. »

Tout comme l'a fait Google DeepMind avec AlphaGo, on peut mettre un réseau de neurones en compétition avec un autre afin de réaliser cet apprentissage. « La partie génération a aussi bien été



En 2016, Google a mis en place un parc de quatorze robots afin de réaliser l'apprentissage d'une I.A. pour lui apprendre à saisir des objets. Il a fallu huit mille heures cumulées aux robots pour réaliser les huit cent mille essais nécessaires.

« Ces derniers mois » ajoute le chercheur. « On met en compétition des réseaux l'un contre l'autre. Nous l'avons fait avec un réseau qui génère des images et le second qui détermine si cette image est réelle ou pas. Les deux réseaux peuvent évoluer ensemble pour générer des images qu'il devient impossible de distinguer du réel. »

LE PASSAGE DU VIRTUEL AU RÉEL POSE PROBLÈME

En robotique, cette approche pose le problème du

couplage entre le modèle simulé et sa mise en exécution dans le robot réel. En effet, dans le monde réel, les capteurs ne fonctionnent pas d'une manière aussi parfaite que dans le simulateur, les gestes et les déplacements peuvent être moins précis, mais surtout les chercheurs doivent faire face à un phénomène plutôt gênant : dans sa simulation, l'intelligence artificielle a cette vilaine tendance à tricher. Cela a été démontré par les chercheurs sur un autre jeu vintage Atari : Q*Bert. L'I.A. créée par les chercheurs de l'uni-

Peut-être arrivera-t-on à faire apprendre à un robot directement

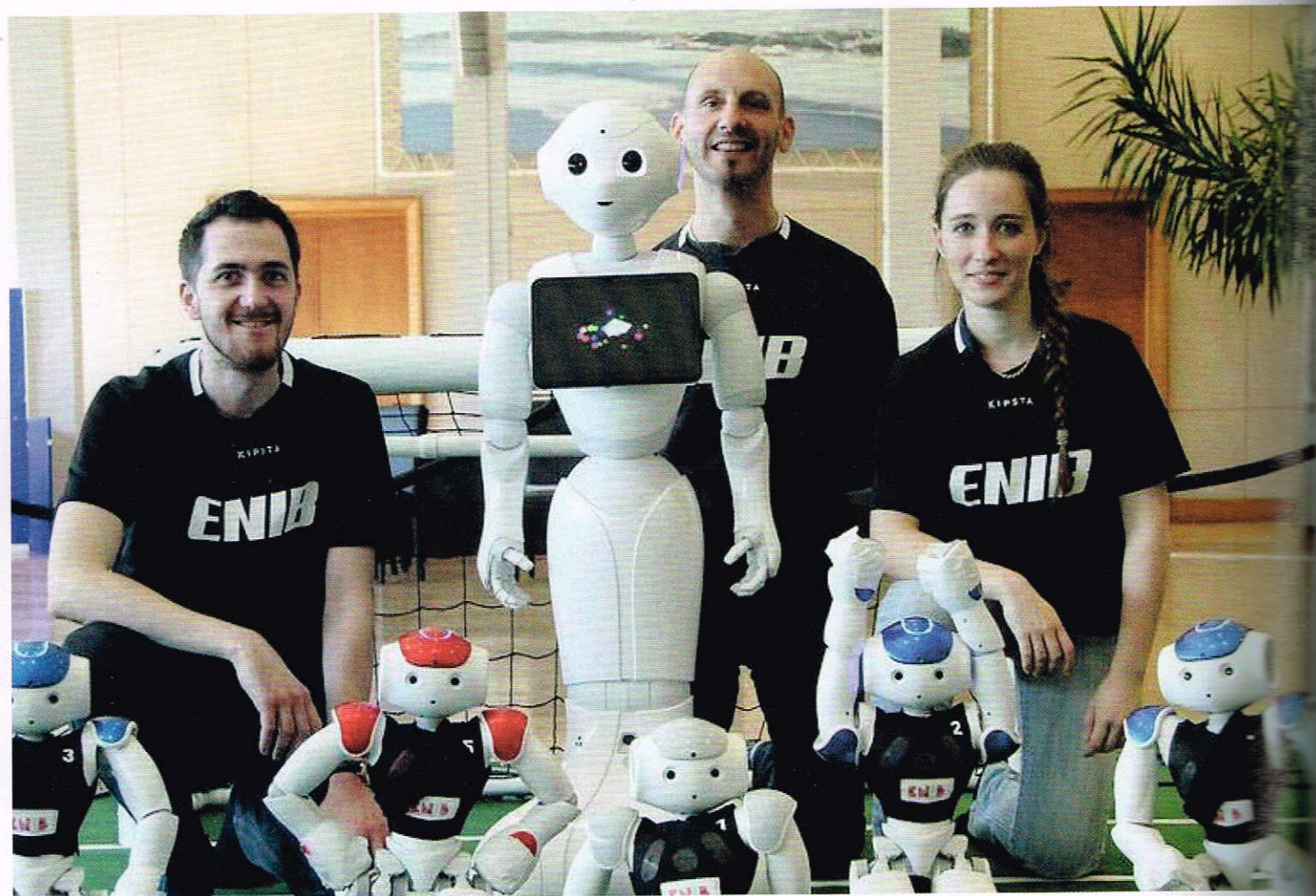


Cédric Buche, chercheur au Lab-STICC du CNRS et responsable du programme transverse intelligence artificielle du laboratoire informatique du Grand Ouest. « Il faut des centaines de milliers d'essais pour obtenir un apprentissage par renforcement

correct, ce qui n'est bien entendu pas envisageable dans le monde réel. On peut encore travailler sur les algorithmes car les temps de convergence sont encore très longs mais de nouveaux algorithmes qui ont des temps d'apprentissage beaucoup plus courts arrivent : avant il fallait des centaines de millions de cycles et qui apprennent maintenant beaucoup plus vite. Peut-être arrivera-t-on un jour à faire apprendre à un robot physique directement en lui faisant observer le geste d'un ouvrier ou un footballeur par exemple. »

versité de Freiburg a trouvé un moyen d'exploiter un bug de ce jeu des années 80 qui lui permet d'amasser les points sans même avoir à passer au deuxième niveau du jeu. C'est une caractéristique typique du *deep reinforcement learning* : comme l'I.A. expérimente des solutions extrêmement diverses avant de converger vers les stratégies les plus efficaces, il va trouver des solutions que jamais un joueur humain ne va tenter car a priori totalement stupides. En l'occurrence, les chercheurs allemands ont utilisé un algorithme évolutionniste qui évolue par petites touches afin de trouver de meilleures solutions. « Comme tous les algorithmes d'apprentissage, ces algorithmes évolutionnistes sont extrêmement opportunistes et vont trouver des caractéristiques de la simulation qui vont leur permettre de maximiser la récompense sans qu'il s'agisse de comportements réalistes » explique Stéphane Doncieux. « On observe ainsi parfois un *reality gap* entre le comportement en simulation et ce que le robot peut faire dans la réalité. Pour contrer cela, nous devons développer des stratégies pour contrer ce biais. »

Avec le passage au monde réel en robotique, ce type de biais n'est plus acceptable puisque le robot ne pourra pas exécuter la solution trouvée par l'I.A. car elle sera physiquement impossible à réaliser. Pour pallier à ce phénomène, il faut améliorer la fidélité des environnements de simulation des robots ainsi que la modélisation du comportement de leurs capteurs, de leurs moteurs et les chercheurs travaillent sur les algorithmes afin de les empêcher de délivrer des résultats aberrants. Autre piste de recherche, améliorer les algorithmes



Dix mille tirs en simulation ont été nécessaires afin d'entraîner le gardien de but de l'équipe Robobreizh.

afin de permettre aux I.A. d'arriver bien plus rapidement à une solution. C'est notamment le thème de recherche de François Lasson, doctorant au LAB-STICC: « Avec le BBN (ou Bayesian Belief Networks), l'apprentissage est rapide et la convergence malgré une faible quantité de données, même si ce

type d'algorithmes est habituellement utilisé sur du big data. Les DB Networks sont des réseaux de croyances et il s'agit d'utiliser des algorithmes génératifs pour générer des comportements. Il fallait faire de l'apprentissage supervisé pour initialement reconnaître des gestes pour les générer et enfin les injecter dans un robot. » Cédric Buche complète cette analyse: « Notre idée a été de placer des marqueurs sur une personne et utiliser un capteur type Kinect pour avoir les données précises sur sa posture puis injecter ces informations dans le réseau de neurones pour détecter le type de gestes, c'est de la classification. Ensuite, on va plus loin avec la génération du geste puis faire exécuter de nouveaux gestes par un robot. » La recherche avance désormais rapidement et ces techniques pourraient bien arriver chez les industriels dans quelques années seulement.

DU DEEP LEARNING POUR MIEUX COMMUNIQUER AVEC LES HUMAINS

Peut-être un jour un robot parviendra-t-il à apprendre simplement en regardant un ouvrier à l'œuvre, mais le deep learning est aussi mis en œuvre afin de faciliter le dialogue entre le robot et les hommes. L'application la plus évidente est le NLP (Natural Language Processing) qui lui permet déjà de saisir les mots d'une phrase. SoftBank Robotics s'est notamment tourné vers Satisfi pour donner la capacité à son Pepper de pouvoir échan-

ger avec un être humain, mais désormais les recherches vont plus en avant car actuellement toute la communication non verbale, c'est-à-dire les gestes, les expressions du visage, les intonations échappent encore aux robots, totalement inconnus si leur interlocuteur est ironique. Le japonais s'est tourné vers un expert de la reconnaissance faciale, EverAI afin de détecter les expressions faciales des interlocuteurs de Pepper. En France, une chercheuse va plus loin encore: Laurence Devillers est professeur d'I.A. à la Sorbonne et chercheuse au CNRS. « Mes recherches portent sur les interactions humain/machine, et notamment sur une partie qui est encore avant-gardiste: comment prendre en compte les émotions des interlocuteurs au cours de l'interaction grâce à des modules de reconnaissance artificielle. Je cherche à savoir comment la Chine peut utiliser ces émotions dans le dialogue, pour, par exemple, générer des réponses qui soient soit empathiques, soit de l'humour. » Elle met en œuvre d'autres stratégies de dialogue. La chercheuse veut embarquer une telle capacité dans des robots afin qu'ils puissent s'apprendre en continu et disposer d'une capacité de converser avec un humain. Le deep reinforcement learning donne les moyens nécessaires aux robots, mais pourrait aussi leur donner un peu plus humains.

“
AVEC LE PASSAGE AU MONDE RÉEL EN ROBOTIQUE, CE TYPE DE BIAIS N'EST PLUS ACCEPTABLE PUISQUE LE ROBOT NE POURRA PAS EXÉCUTER LA SOLUTION TROUVÉE PAR L'I.A. CAR ELLE SERA PHYSIQUEMENT IMPOSSIBLE À RÉALISER.

”