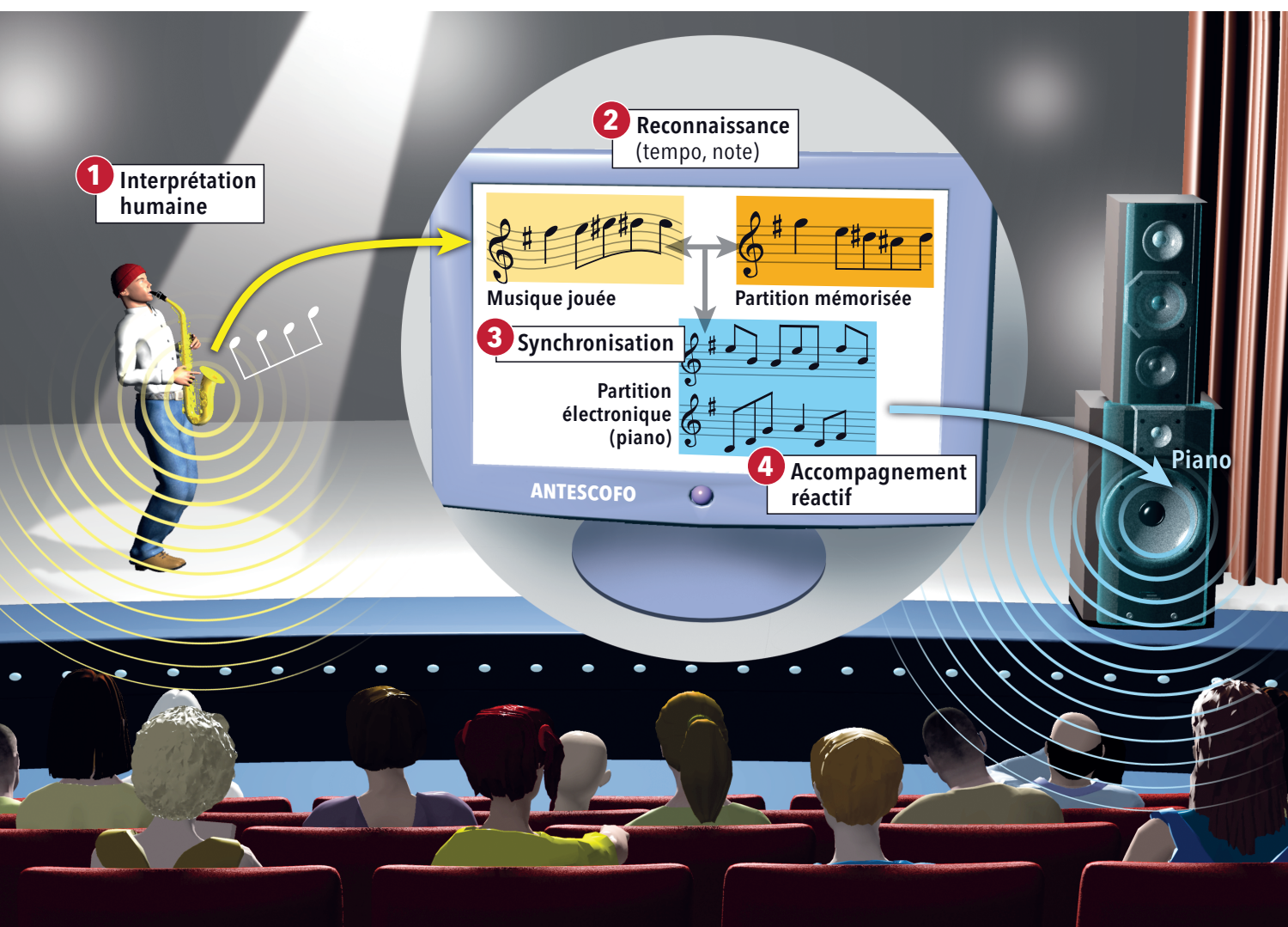


L'ordinateur qui joue comme un musicien

Doté de la capacité d'interaction propre à l'homme, le logiciel Antescofo est un musicien à part entière. Il est capable d'interpréter une partition tout en s'adaptant au jeu des instrumentistes.



Le logiciel Antescofo enregistre les sons produits par un musicien humain (1). Il analyse ensuite la musique jouée et la compare avec la partition qu'est censée suivre l'interprète, et qui est mémorisée dans l'ordinateur (2). Antescofo peut non seulement reconnaître les notes jouées, mais aussi suivre le tempo de l'interprétation. Ces informations sont alors synchronisées (3) avec le programme informatique chargé de jouer la partition d'accompagnement. L'ordinateur adapte ainsi son jeu en temps réel : il ralentit ou accélère en réaction au rythme imposé par le musicien (4).



PAR **Arshia Cont**,
chercheur à l'Ircam et responsable
du projet MuSync.

LAURÉAT du Prix La Recherche
2011 mention coup de cœur.

Le 20 février 2012, le Français Philippe Manoury a reçu la Victoire de la musique classique du meilleur compositeur, pour son opéra *La Nuit de Gutenberg*. Cette récompense, il la doit bien sûr à son talent, mais aussi à un musicien pas comme les autres : le logiciel Antescofo. En effet, ce programme informatique, conçu par l'Institut de recherche et coordination acoustique/musique (Ircam) est capable de jouer une partition de musique via la carte son d'un ordinateur. Mais surtout, il modifie son jeu en suivant celui des autres interprètes, à l'image d'un musicien dans un ensemble.

L'idée d'Antescofo a vu le jour en 2007 sous l'impulsion du compositeur italien Marco Stroppa. À l'occasion de la commande d'une œuvre pour saxophone et électronique au Japon, il s'était lancé un défi à la fois artistique et scientifique. Il souhaitait faire interagir un musicien et un ordinateur sur scène pendant le concert, comme si ce dernier était humain.

Doter une machine de cette capacité d'interaction est complexe : il mobilise des techniques issues des recherches les plus avancées en traitement du signal, apprentissage automatique, intelligence artificielle et programmation des systèmes temps réel. Face au défi technologique, l'équipe Interactions musicales temps réel de l'Ircam et moi-même avons été associés au projet.

Une première étape a consisté à étudier le vivant pour mieux comprendre la nature de la coordination musicale

chez l'homme. Les musiciens forment en effet un groupe parfois composé de centaines d'individus sur scène. Chacun est responsable d'interpréter sa propre partition en cohérence avec celle des autres : il doit en particulier ralentir ou accélérer pour se synchroniser avec ses collègues. Le résultat final est alors très proche de celui que le compositeur avait imaginé et transcrit sur des partitions de musique.

Écouter et agir. Cette aptitude à se coordonner musicalement malgré toutes les variations liées à une interprétation est l'une des compétences de base des musiciens. Mais comment l'ordinateur peut-il s'intégrer à cette microsociété présente sur scène, et ainsi devenir un musicien à part entière ? Pour parvenir à un tel résultat, il doit être capable de mener deux tâches en parallèle : écouter et agir. Afin d'interagir avec son environnement, une machine doit en effet percevoir, sur le vif, ce qui l'entoure. Elle doit être capable d'« entendre » le jeu des interprètes en temps réel. Dans ce but, un système d'écoute a été mis au point afin de suivre pas à pas les sons émis par les instrumentistes. Il est constitué de deux logiciels.

Le premier, baptisé « agent audio », est un programme qui a mémorisé une partition et cherche à la reconnaître lorsqu'elle est jouée, à partir des sons captés par un microphone. Il repose

***LA DYNAMIQUE** désigne la manière dont est jouée la note. Celle-ci peut être plus ou moins accentuée, plus ou moins maintenue ou plus ou moins étouffée par exemple.

sur un système de reconnaissance probabiliste, similaire aux dispositifs de suivi des missiles ou aux systèmes de reconnaissance vocale.

Son principe de fonctionnement peut-être comparé au système de suggestion automatique de mots utilisé lors de la saisie d'un SMS. Dans le cas du téléphone, une liste de mots probables est proposée en fonction des lettres préalablement inscrites. De son côté, Antescofo fait plusieurs hypothèses sur la note à venir en fonction de la partition mémorisée. Plus le temps passe, plus la note suivante commence à se « faire entendre », et plus la liste des hypothèses se réduit jusqu'à aboutir à la hauteur de la note réellement jouée. La différence, ici, est que l'événement doit être reconnu par Antescofo dès son arrivé, à la volée, ceci pour permettre un temps de réaction instantané par le système. En effet, des variations peuvent toujours intervenir lors d'une exécution en direct.

Seul, l'agent audio ne suffirait pas à obtenir une écoute performante. En effet, les sons à détecter ne sont pas tous consignés dans la partition. La notation baroque en est un bon exemple : l'écriture comporte certains éléments qualifiés d'absolus, comme la hauteur de la note, et d'autres qui sont relatifs, c'est-à-dire laissés à la discrétion de l'interprète à l'intérieur de certaines limites, comme le tempo ou les dynamiques* par exemple.

De même, la variabilité introduite par les différents types d'instruments et les techniques de jeu propre à chaque musicien doivent être prises en compte. Tel pianiste appuiera particulièrement fort sur les touches de son instrument avant de jouer un accord, tel violoniste tiendra son archet d'une façon différente de celle d'un autre, tel flûtiste produira un staccato plus court que ceux de ses collègues.

Il ne faut pas non plus oublier la difficulté de reconnaissance dans un environnement soumis aux bruits de >>>

L'essentiel

> **LE LOGICIEL ANTESCOFO** interprète une partition de musique tout en s'adaptant au jeu des instrumentistes.

> **DOTÉ D'UN SYSTÈME** d'écoute interactif, il ralentit ou accélère en réaction au rythme imposé par les autres musiciens.

> **LES SYSTÈMES D'ÉCOUTE** et d'accompagnement sont synchronisés en temps réel.

L'ordinateur qui joue comme un musicien

» fond, aux différences d'accordage et aux biais des dispositifs de captation. Or, ce sont toutes ces fluctuations qui donnent à chaque interprétation son caractère unique. C'est en ce sens que toute partition destinée à un interprète est dite « virtuelle » : elle ouvre des champs de possibles, sans les déterminer complètement.

Confrontés à ces nombreux obstacles, nous avons fait appel à un autre partenaire. Nous avons sollicité Claude Delangle, professeur au Conservatoire national supérieur de musique et de danse de Paris, pour étudier de près ces phénomènes aléatoires. Et ce, pour tout style de musique confondu, de Maurice Ravel à Astor Piazzolla en passant par Steve Reich et la création contemporaine européenne. Ses observations ont permis de recenser les différents types d'interprétations possibles pour une même œuvre. Ce qui a ensuite permis d'identifier les paramètres jouant un rôle clé dans ces variations. Parmi ceux-ci : le tempo.

Cette analyse conduite par Claude Delangle a débouché sur la mise au point du second logiciel du système d'écoute, baptisé « agent temporel ».

Le défi polyphonique

Une des méthodes les plus efficaces pour reconnaître la partition jouée par un instrument consiste à analyser les signaux acoustiques captés par un simple micro. Le temps de réaction est alors bien plus rapide que celui d'une oreille humaine exercée. Toutefois, cette technique a ses limites. Elle ne permet pas d'identifier deux sons superposés. « Une polyphonie, même minimale, semble pour l'instant hors des possibilités actuelles de détection audio, car deux sons qui se superposent, mélangent tous leurs harmoniques et l'on ne sait plus à quelles fondamentales chacun appartient »,

Son rôle : synchroniser la machine et le musicien, pour s'adapter aux variations de tempo. On peut comparer son principe de fonctionnement à la « sympathie des horloges », un phénomène de physique découvert au XVII^e siècle par Christian Huygens. Ce dernier avait montré que deux pendules oscillants finissaient par se balancer au même rythme s'ils étaient reliés avec des lames de profil correct. Dans le cas d'Antescofo, l'agent temporel joue le rôle de la lame. À tout moment, l'horloge interne de l'ordinateur est ainsi couplé au rythme du musicien.

Compétition et collaboration. Le système d'écoute « interactif » d'Antescofo est aujourd'hui composé de ces deux agents. Ils sont en collaboration et en compétition permanente : l'agent temporel aide l'agent audio à anticiper de futurs événements (notes, dynamiques), et l'agent audio aide l'agent temporel à décoder la bonne vitesse (tempo). Le résultat final du décodage se fait donc en fonction de la certitude soit temporelle soit événementielle de chaque agent. Comme les êtres humains, Antescofo combine donc plusieurs sour-

explique Philippe Manoury, compositeur français. Cela pose problème. En effet, lors un concert, flûtes, hautbois, clarinettes, pianos, guitares et autres instruments se côtoient. Il faut donc réussir à reconnaître autant de partitions qu'il y a d'instruments. Pour y parvenir, Antescofo ne s'appuie pas que sur la reconnaissance des hauteurs. Il dispose dans sa base de données d'un ensemble de profils sonores, sortes de carte d'identité, décrivant le mode de fonctionnement de chaque instrument de l'orchestre. Ce qui lui permet de distinguer le son d'une flûte par rapport au son d'un piano.

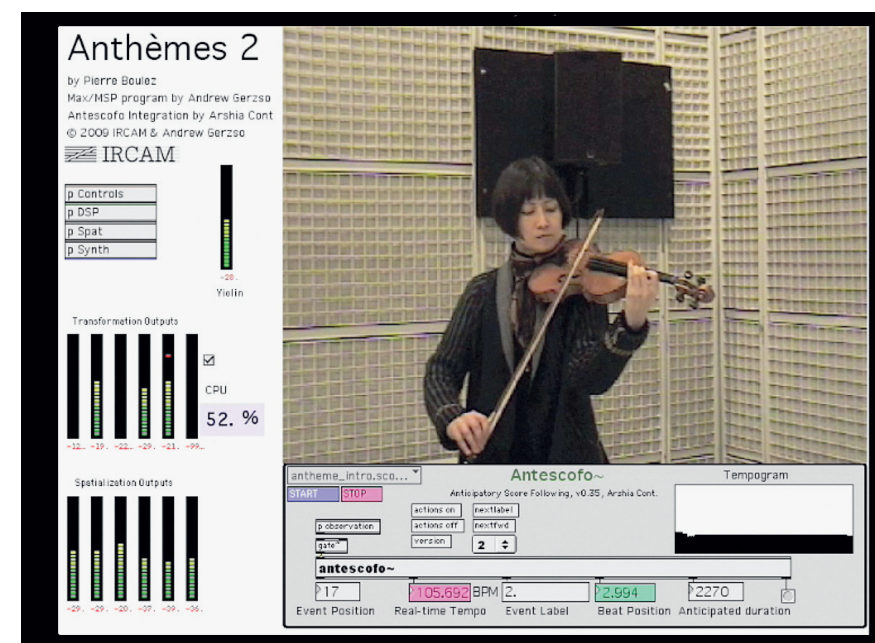
ces d'information, chacune faillible, pour arriver à une écoute fiable.

Mais Antescofo ne se contente pas d'écouter. Il permet aussi à l'ordinateur d'agir, c'est-à-dire de jouer sa propre partition électronique, en réaction au jeu de l'instrumentiste. Cet « accompagnement réactif » nécessite toutefois une certaine capacité d'anticipation. En effet, pour pouvoir jouer ensemble, les musiciens humains, en plus d'une écoute élaborée, anticipent le jeu de leurs collègues afin de se synchroniser avec eux. Cela concerne autant la temporalité du jeu que les dynamiques de l'interprétation.

À cet égard, il ne faut pas parler du temps mais des temps musicaux. En effet, une partition de musique occidentale classique contient des événements obéissant à des échelles de temps différentes qui se superposent. Les opéras débutent par exemple par une ouverture orchestrale décomposée en plusieurs mouvements, qui précède plusieurs actes. Ce temps « macroscopique » diffère du temps « microscopique » qui désigne la durée de chacune des notes. Dans l'intervalle de ces temps extrêmes on trouve d'autres échelles, celles de la mesure, du rythme, de l'harmonie, de la mélodie.... De plus, l'écriture musicale, à travers les siècles, a permis aux compositeurs d'écrire ces temps en valeurs relatives, donnant une certaine liberté au musicien au moment de l'interprétation.

L'existence de ces différentes échelles de temps représente un défi majeur pour l'informatique musicale, à la fois pour son écriture et pour son exécution en temps réel. Pour répondre à cet enjeu, Antescofo a été doté d'un langage de coordination assez similaire aux langages dédiés à la programmation des systèmes temps réels*, comme la commande de vol des Airbus par exemple. Son rôle est de lancer les programmes musicaux aux moments indiqués en réaction du jeu des musiciens, et de les synchroniser entre eux et avec les interprètes.

*UN SYSTÈME est en temps réel lorsque ses actions logicielles, suffisamment rapides, dirigent des processus en cours d'exécution.



Une violoniste joue *Anthèmes 2* du compositeur Pierre Boulez, accompagnée du logiciel Antescofo. Celui-ci joue sa propre partition électronique et réagit en temps réel aux variations de l'instrumentiste, en modifiant son tempo notamment. © IRCAM

Dans ces langages, qualifiés de synchrones, on déclenche une action simultanément à la détection d'un événement, c'est-à-dire avec un temps de réaction quasi nul. Bien sûr, dans la machine, l'action est réalisée après la survenue de l'événement, mais on maîtrise ce décalage afin qu'il devienne négligeable. Pour Antescofo, le temps de réaction est ainsi inférieur à 15 millisecondes, ce qui le rend imperceptible.

Antescofo a joué en public

- *Speakings*, de Jonathan Harvey (Orchestre écossais de la BBC, 2008)
- *Explosante-Fixe*, de Pierre Boulez, (Orchestre philharmonique de Los Angeles, 2010)
- *Einspielung*, d'Emmanuel Nunes (Festival Agora, 2011)
- *La Nuit de Gutenberg*, de Philippe Manoury. (Opéra national du Rhin, 2011)
- *Re Orso*, un opéra de Marco Stroppa (Opéra comique de Paris, 2012)

Le couplage du système d'écoute interactive et d'accompagnement réactif est l'idée principale derrière le projet Antescofo. Une première version du logiciel a vu le jour en 2009. Capable de se tenir aux côtés des musiciens pour participer à l'interprétation de l'œuvre, elle ouvre la voie à de nouveaux paradigmes musicaux. Une application immédiate d'un tel système est l'accompagnement automatique d'un interprète solo. Un saxophoniste peut se faire accompagner d'un piano virtuel par exemple. L'ordinateur joue alors le rôle d'un musicien et réagit, en temps réel, au jeu de l'instrumentiste.

Le langage synchrone d'Antescofo permet d'ajouter, aux événements musicaux d'un accompagnement classique, d'autres programmes informatiques multimédias. On peut notamment intégrer des effets sonores, contrôler la lumière et diffuser des vidéos. Autant d'actions supplémentaires destinées à enrichir le vocabulaire musical d'éléments scéniques. Parmi les compositeurs intégrant ces technologies dans leurs œuvres, le premier à avoir exploré

ces possibilités est Philippe Manoury. Son opéra *La Nuit de Gutenberg* est composé d'acteurs, de chanteurs, d'un chœur, d'un orchestre, de vidéos et de dispositifs électroniques. Cette mise en présence du vivant et du numérique devient une interprétation cohérente, jouée avec Antescofo en partie.

Orchestres prestigieux. Depuis sa création, Antescofo a fait l'objet de tests et d'améliorations grâce à une collaboration fertile avec plusieurs compositeurs dont Jonathan Harvey, Philippe Manoury, Emmanuel Nunes et Pierre Boulez. Il est aujourd'hui utilisé lors de concerts publics avec divers artistes et ensembles dont l'Orchestre de Radio France, l'Orchestre écossais de la BBC, l'Orchestre philharmonique de Berlin, l'Orchestre philharmonique de Los Angeles et des salles d'opéra. Son répertoire consiste en plus de trente œuvres.

L'aventure ne s'arrête pas là. L'écriture musicale, avec toute sa richesse d'expression et sa réalisation scénique, pose toujours des défis à la conception de langages informatiques dédiés ainsi qu'au domaine de l'intelligence artificielle. Ce qui fait l'originalité du musicien humain lors d'une exécution est un point essentiel qui nécessitera beaucoup d'avancées pour pouvoir l'intégrer dans l'intelligence des machines. L'idée globale est de doter les ordinateurs d'un langage aussi riche que celui des humains. Dans ce but, Antescofo mobilise désormais des chercheurs, des artistes et des interprètes au sein d'une équipe-projet commune avec l'Ircam, Inria et le CNRS, baptisée MuSync. ■

Pour en savoir plus

- <http://repmus.ircam.fr/antescofo> Le site Web du projet Antescofo.
- www.youtube.com/watch?v=iN9MmiSIBj0 Une démonstration d'accompagnement automatique avec Antescofo.

PROCHAIN RENDEZ-VOUS MUSICAL AVEC ANTESCOFO

- Le 1^{er} juin 2012 à la Salle Pleyel (Paris), dans un concerto pour piano, orchestre et électronique, composé par Philippe Manoury, avec l'Orchestre de Paris.