The Muddy Children Puzzle
Stage sur la Logique Modale et Épistémique

Gatien HADDAD

L2 CMI-Info 2020-2021

Maître de stage : Christian RETORE

Juin 2021

Table des matières

1 Introduction et contexte

La logique épistémique est une logique permettant de raisonner sur les connaissances d'un individu, appelé agent ou acteur, et également sur les connaissances au sein d'un groupe et entre les acteurs. Elle introduit une modalité pour chaque agent selon ses connaissances. C'est Jaakko Hintikka et Edward John Lemmon qui ont introduit la logique épistémique au grand public au travers de divers ouvrages parus vers la fin des années 50 [1]. La logique épistémique est une branche de la logique modale, cette dernière apportant des informations sur la « qualité du vrai », au travers de quatre modalités : la nécessité, l'impossibilité, la possibilité et la contingence.

Le but de ce stage était d'être une introduction à la logique épistémique, de comprendre la modélisation de problèmes de logique épistémique et le concept de « mondes possibles ». Une fois à l'aise, l'objectif était de trouver des exemples, des puzzles, des énigmes, et de les résoudre grâce à une approche mathématique, en appliquer les connaissances acquises et tenter d'expliquer ces mêmes exemples à un néophyte dans le domaine.

Nous allons dans un premier temps voir la logique épistémique à travers l'exemple du Muddy Children Puzzle (ou "puzzle des enfants sales"). Dans ce puzzle, des enfants ont, ou n'ont pas, de la boue sur le front, mais ils ne peuvent pas se voir eux même. En voyant le front de leurs frères et sœurs, ils vont à un moment donné être capables de déterminer s'ils ont ou non de la boue sur le front.

Dans un second temps, nous allons expliquer la syntaxe de manière plus fondamentale, à l'aide de la notion de savoir et de croyance, et également expliquer plus en détail les modèles de Kripke. Nous verrons ensuite deux exemples plus poussés pour développer les connaissances acquises et tenter de les expliquer.

Le premier exemple étant une énigme répandue avec quatre prisonniers et des chapeaux sur leur tête : pour gagner la liberté, ils devront deviner la couleur du chapeau sur leur tête sans pouvoir le voir.

Le second est un jeu de cartes se jouant à trois joueurs, où chaque joueur met ses deux cartes sur la tête de sorte à ce qu'il ne voit pas ses cartes, mais puisse voir celles des autres. Le but est de deviner les cartes que l'on a sur la tête en fonction de celles des autres.

Je tiens à remercier M. Christian Rétoré pour m'avoir guidé durant toute la durée du stage. Il a été à la fois disponible et pédagogue, et m'a conseillé des ouvrages et références très utiles à la compréhension de la logique épistémique et modale. J'aimerais également remercier Davide Catta (doctorant LIRMM), avec qui j'ai eu plusieurs fois l'occasion d'échanger et qui m'a éclairé sur certains sujets.

2 Du Muddy Children Puzzle à la logique épistémique

L'énoncé du puzzle des enfants sales est le suivant [2] [9] : Un père laisse ses enfants aller jouer dans le jardin, malheureusement, il a plu la veille, le père demande alors aux enfants de ne pas se salir dans la boue. Évidemment, chaque enfant évite au plus possible de se tacher, mais prendrait un malin plaisir à voir ses frères et sœurs se faire gronder s'ils venaient à se salir.

Cependant, les enfants qui se salissent ne le font que sur le front, Un enfant sale ne peut donc pas voir son front pour savoir s'il est taché, mais peut cependant voir celui des autres. Lorsqu'ils rentrent à la maison, aucun enfant n'a communiqué avec ses frères et sœurs ; faisant face à leur père mécontent, ils restent silencieux. Ce dernier s'exclame alors :

- "Au moins l'un d'entre vous est sale.", puis il ajoute
- "Si l'un d'entre vous sait qu'il est sale, qu'il lève la main.". Le père répète alors cette phrase tant qu'aucun enfant n'a levé la main.

Le but de ce puzzle est de comprendre et de prouver que, pour un nombre k d'enfants sales, ils lèveront tous la main en même temps au bout de la k-ième fois que le père répètera sa phrase. Comme beaucoup de problèmes de ce type, on peut réduire le problème à un petit nombre d'agents pour un raisonnement plus intuitif.

2.1 Cas avec 2 enfants

Imaginons qu'il n'y ait que deux enfants : Alice et Bob. Alice voit donc le front de Bob, et Bob voit le front d'Alice, au moment où les deux rentrent dans la pièce, on va pour l'instant considérer qu'ils sont tous les deux sales.

Pour ce premier cas, le problème peut se résoudre de manière assez intuitive ; une fois que le père a dit sa première phrase (qui est donc "au moins l'un d'entre vous a de la boue sur le front"), Alice et Bob se regardent. Alice voit de la boue sur Bob, et sait qu'il y a au moins l'un des deux qui est sale. Elle hésite donc entre le cas où "seul Bob a de la boue sur le front", et "Alice et Bob ont de la boue sur le front". Bob, lui, hésite entre "seule Alice est sale" et "nous sommes tous les deux sales".

Ces cas sont appelés des *mondes possibles*, et on peut déjà remarquer qu'Alice et Bob envisagent un monde possible commun (celui où ils sont tous deux sales), et c'est également le seul. Cependant, ils ne savent pas ce que voit l'autre, et donc ce qu'il sait. Ils n'ont donc, pour l'instant, aucun moyen de trancher entre les deux mondes qu'ils pensent possible. On peut représenter ces trois mondes par le schéma suivant [Figure 1, p5] [9].

La notation est la suivante ; As et Bs signifient que respectivement Alice et Bob sont sales.
Le symbole $\vdash$ signifie « ce que force ce monde », ou autrement dit, ce que l'on peut en déduire. Par exemple « 1 $\vdash$ As » signifie que l'on peut déduire du monde 1 que Alice est sale, ou que le monde 1 force le fait qu'Alice soit sale (voir page 6 pour les flèches).

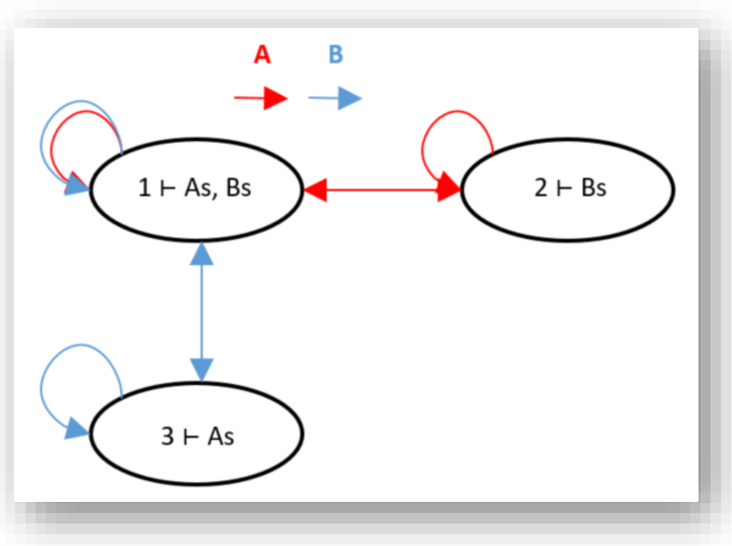

Figure 1 - Les trois mondes considérés possibles pour Alice et Bob

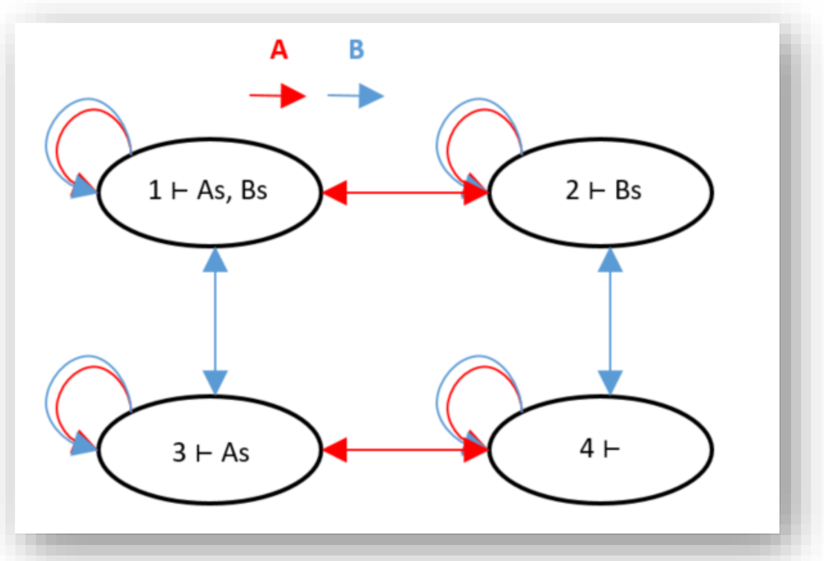

Figure 2 - Tous les mondes possibles pour Alice et Bob

Les flèches sur le schéma, rouges pour Alice et bleues pour Bob, représentent la considération. Cela signifie que, dans le cas du monde 1, comme dit préalablement, Alice considère possible le monde 2 et le monde 1, mais également qu'elle ne peut pas les différencier. Même cas pour Bob, qui depuis le monde 1 considère le monde 1 ainsi que le monde 3, et qui ne peut donc pas faire la différence entre ces deux mondes. Pour la suite, on considèrera que tous les diagrammes sont réflexifs et on omettra les boucles.

Le père va alors poser la question pour savoir si un enfant sait ou pas s'il a de la boue sur le front. Alice et Bob restent donc stoïques car ils ne peuvent pas être sur du monde dans lequel ils se trouvent.

Mais paradoxalement, c'est en ne disant rien qu'ils vont transmettre des informations. Pour Alice, le fait que Bob n'ait pas levé la main lui indique que Bob ne sait pas si il est sale ou non, donc que Bob hésite entre plusieurs mondes distincts. Alice comprend donc que Bob hésite entre le monde 1, où lui et Alice sont sales, et le monde 3 où seule Alice est sale. Or, dans ses deux mondes, Alice est sale ; elle en déduit donc que du point de vue de Bob, elle est forcément sale.

Évidemment, Bob fait le raisonnement identique, et conclut qu'il est lui-même également sale. Lorsque le père va donc répéter une seconde fois sa phrase pour savoir si un enfant sait qu'il est sale, Alice et Bob vont tous deux lever la main.

Si on souhaite généraliser, au moment où Alice et Bob entrent dans la pièce, il y a au plus quatre configurations possibles :

- Alice est propre et Bob est sale
- Alice est sale et Bob est propre
- Ils sont tous les deux sales
- Ils sont tous les deux propres.

Tous ces mondes possibles sont représentés par le schéma ci-dessous [Figure 2, p5]. Et c'est au moment où le père déclare qu'il y a au moins un enfant sale que le monde 4 disparaît. En effet, le monde 4 est celui où aucun enfant n'est sale, ce qui entre en contradiction avec ce que vient de dire le père ; Alice et Bob ne considèrent donc pas ce monde comme possible. La disparition de ce monde entraîne également les flèches qui y menaient, car ces dernières signifiaient simplement que ce monde était considéré par Alice ou Bob.

2.2 Cas avec 3 enfants

C'est au tour d'un troisième enfant, Chloé, de venir se joindre aux deux autres [9]. Le total des mondes possible est alors doublé, car pour chaque monde impliquant auparavant Alice et Bob, il en existe un autre avec Chloé. On a alors un schéma montrant les 8 mondes possibles, cependant comme précédemment, une fois que le père a dit qu'il y a au moins un enfant de sale, le monde 4 n'a plus lieu d'être, et il est donc retiré des mondes possibles [Figure 3, p7] (pour rappel, les schémas sont toujours réflexifs).

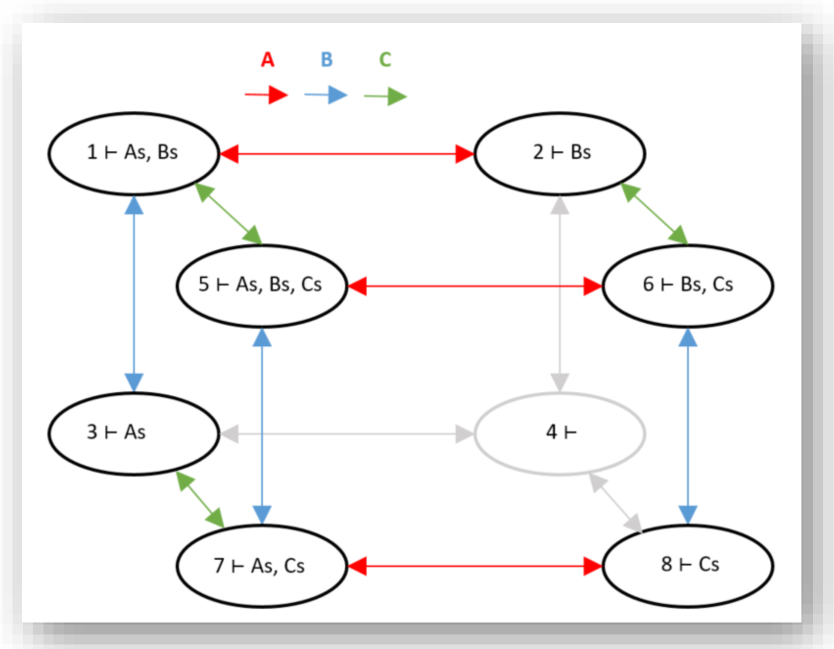

Figure 3 - Les huit mondes possibles pour Alice, Bob et Chloé.

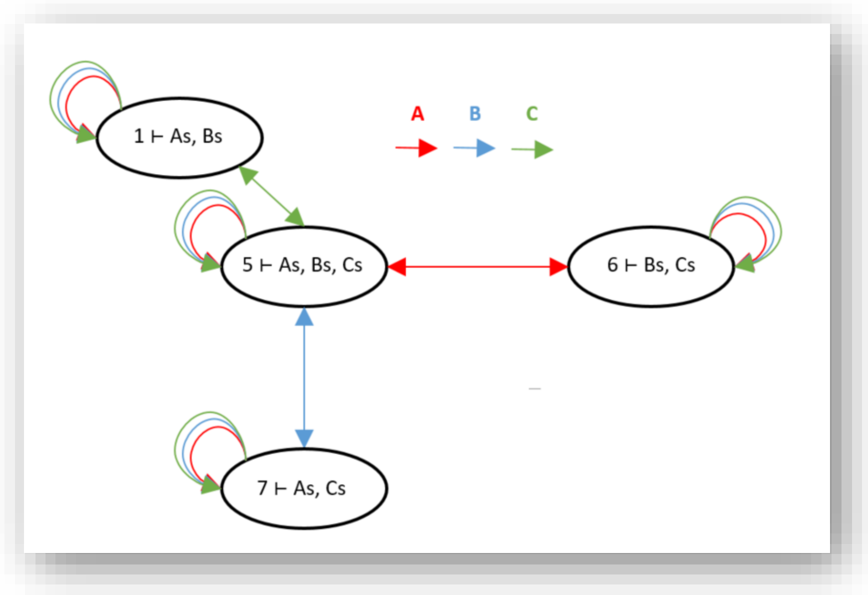

Figure 4 - Les 4 mondes possibles restant.

Nous allons pour cet exemple, considérer que nous nous trouvons dans le monde 5, à savoir que tous les enfants sont sales. On constate que dès que les enfants entrent, ils peuvent voir de la boue sur le front de leurs deux frères et sœurs.

Une fois de plus, le père prononce la phrase “Au moins l’un d’entre vous a de la boue sur le front”, puis “si l’un de vous sait qu’il est sale, qu’il lève la main.”. Comme chaque enfant en voit deux autres avec de la boue sur le front, aucun ne réagit, cela donne l’information commune qu’il y a au moins deux enfants avec de la boue sur le front. En effet, si Alice était par exemple la seule sale, elle aurait vu le front de Bob et de Chloé propre, et se serait donc déclarée immédiatement. On peut donc éliminer les mondes 2, 3 et 8, qui impliquaient respectivement Bs, As et Cs. On se retrouve alors avec 4 mondes possibles [Figure 4, p7].

On a donc que Alice hésite entre les mondes 5 et 6, où elle sait que Bob et Chloé sont sales ; Bob lui hésite entre les mondes 5 et 7, où il sait que Alice et Chloé sont sales ; et pour Chloé, c’est entre les mondes 1 et 5, sachant Alice et Bob tous deux sales.

Le père repose alors la question “Si l’un de vous sait s'il est sale, qu’il lève la main”. Encore une fois, les trois enfants ne répondent pas. Mais cette fois encore, le fait de ne pas parler apporte son lot d'informations : du point de vue d’Alice, elle sait que Bob et Chloé sont sales, mais elle vient également d’apprendre qu’aucun des deux *ne sait* s’il est sale.

Or comme on peut le voir pour Bob sur le schéma des mondes restants, Bob ne peut pas se trouver dans le monde 6, car si c’était le cas, il saurait qu’il est sale (car il n’y a qu’une flèche bleue qui part du monde 6 et qui y retourne, donc Bob ne considèrerait que ce monde), et il se serait donc déclaré avec Chloé, qui aurait suivi le même raisonnement. Mais Bob ne s’est pas déclaré, et comme Alice sait que ça ne peut pas être les mondes 1 et 7, car elle voit Bob et Chloé couverts de boue, elle peut donc conclure que c’est obligatoirement le monde 5 qui est le seul monde possible restant. Chloé et Bob, à travers un raisonnement identique, la rejoignent, et lèvent tous les trois la main lorsque le père pose la question pour la troisième fois.

2.3 Généralisation pour n enfants

Nous avons vu que lorsqu'il y a deux enfants sales, ils peuvent tous les deux se déclarer à la seconde question du père, et pour trois enfants, à la troisième question. On peut alors intuitivement se dire que lorsqu’il y a un nombre n d’enfants sales, ou disons, un nombre k d’enfants sales parmi n enfants, tous les enfants sales pourront alors se déclarer au bout de la k-ième fois que le père répètera sa phrase.

De manière générale, lorsque le père demande si les enfants savent s’ils sont sales pour la x-ième fois et que ces derniers ne répondent pas, cela revient à dire « Tous les enfants ne savent pas s’ils sont sales » et également « Il y a au moins x+1 enfants avec de la boue sur le front ».

3 Formalisation de la logique modale et modèles de Kripke

Nous allons à présent voir les définitions plus formelles et la syntaxe de la logique modale et de la logique épistémique qui viendront en complément des connecteurs de base de la logique classique propositionnelle.

Pour rappel, les 5 connecteurs de la logique classique sont : $\neg$ (négation), $\wedge$ (conjonction), $\vee$ (disjonction), $\rightarrow$ (implication) et $\leftrightarrow$ (équivalence). Il y a de plus les symboles propositionnels p, q, r... pouvant prendre les valeurs « vrai » ou « faux » (respectivement 1 et 0). Viennent également s'ajouter $\top$ (Top, toujours vrai), et $\bot$ (Bottom, toujours faux).

3.1 Logique Modale

La logique modale a pour but de venir « compléter » la logique propositionnelle, en introduisant deux *modes* ou opérateurs de modalités, que l'on peut également appeler des opérateurs unaires. Le premier, $\square$, est noté carré, et le second, $\Diamond$, se lit losange [5].

- $\square(\varphi)$ signifie alors « il est *nécessaire* que φ »,
- et $\Diamond(\varphi)$ signifie « il est *possible* que φ ».

On peut également utiliser les autres connecteurs de la logique propositionnelle pour former d'autres formules, telles que $\square(\neg\varphi)$, signifiant « Il est nécessaire que non φ », ou $\neg\Diamond(\varphi)$, pour « Il n'est pas possible que φ » [3][6].

On peut étendre l'utilisation de carré et losange pour exprimer les connaissances des agents. On notera alors [i] pour « l'agent i sait que », et <i> pour « l'agent i croit possible que ». On peut par exemple écrire <B>(«Il pleut») pour « Bob croit possible qu'il pleuve », ou encore [A](<B>(«Il pleut»)) pour « Alice sait que Bob croit possible qu'il pleuve » [8].

Il existe également, en logique épistémique la modalité K, tel que « $K_i\ \varphi$ » signifie « l'agent i sait que (Knows) φ » ; mais sémantiquement équivalente à [i], nous garderons cette dernière [1].

Il y a également la notion de « forcing », notée $\Vdash$, où $M \Vdash \varphi$ signifie que le monde M force φ. On a les équivalences suivantes :

- « $M \Vdash \varphi \wedge \psi$ » $\leftrightarrow$ « $(M \Vdash \varphi) \wedge (M \Vdash \psi)$ » (marche également avec la disjonction).
- « $M \Vdash [i]\varphi$ » si et seulement si « pour tout monde N en relation avec M, on a $N \Vdash \varphi$ ».

Comme dit précédemment, la notion de forcing permet de rendre compte de ce qu'il est déductible d'un monde, ou ce qu'être dans ce monde implique comme propositions.

3.2 Systèmes d'axiomes

En logique modale, il existe également des systèmes composés d'axiomes [1] :

Le système K, constitué des axiomes

- (K) $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$, (ou avec des agents $[i](\varphi \rightarrow \psi) \rightarrow ([i]\varphi \rightarrow [i]\psi)$) :
  « Si l'agent i sait que φ implique ψ, alors cela implique que, si l'agent i sait que φ, alors il sait que ψ », ou φ et ψ, deux propositions (Axiomes de distribution de Kripke).
- (NR) « Si φ est un théorème dans K (donc que la proposition φ est toujours vraie), alors $\Box\varphi$ (ou $[i]\varphi$) » (Règle de nécessité).

Pour former le système M (parfois appelé « T »), on rajoute l'axiome (M) au système K :

- (M) $\Box\varphi \rightarrow \varphi$ (ou $[i]\varphi \rightarrow \varphi$) : « Si l'agent i sait que φ, alors φ est valide (toujours vraie).

Viennent à cela s'ajouter les systèmes S4 et S5, qui se constituent des axiomes suivants :

- (4) $\Box\varphi \rightarrow \Box\Box\varphi$ « Si l'agent i sait que φ, alors *il sait qu'il sait* que φ »
- (5) $\Box\varphi \rightarrow \Box\Diamond\varphi$ « Si l'agent i sait que φ, alors il sait qu'il croit possible que φ ».

Le système S4 se construit en ajoutant (4) au système M, de même pour S5 qui se construit en ajoutant (5) au système M.

L'intérêt de ces deux systèmes est qu'ils permettent la simplification de formules, en évitant les itérations superflues ; avec S4 on a les égalités :

- $\Box\Box\ldots\Box\varphi = \Box\varphi$
- $\Diamond\Diamond\ldots\Diamond\varphi = \Diamond\varphi$

Et pour S5 on a :

- $**\ldots\Box\varphi = \Box\varphi$
- $**\ldots\Diamond\varphi = \Diamond\varphi$

Où les * peuvent à la fois représenter carré ou losange. Cela peut paraitre contre intuitif au premier abord qu'une formule comme $\Diamond\Diamond\Box\varphi$ dans S5 puisse être simplifié par $\Box\varphi$, mais cela permet de raccourcir la taille des formules et donc de les comprendre plus aisément [5].

3.3 Modèles de Kripke

Les modèles utilisés précédemment pour représenter les différents mondes sont appelés des modèles de Kripke. Saul Kripke est un logicien et philosophe Américain, il est considéré comme un pionnier de la logique modale grâce à des articles qu'il a publiés dans les années 60 [7]. Les modèles de Kripke servent à représenter les différents mondes possibles et leurs interactions.

La notion de flèches que nous avons vus auparavant est donc une relation d'accessibilité entre les différents mondes. Comme les modalités carré et losange donnent une « portée » aux propositions (qui peut être plus ou moins grande), la relation d'accessibilité permet de visualiser les mondes où une proposition est valide. En langage formel, « M R_i N » signifie que le monde M et N sont en relation par l'agent i, et donc que ce dernier ne peut pas les différencier. Comme R est une relation d'équivalence, on sait qu'elle est donc réflexive, symétrique et transitive. Comme nous l'avons précisé plus haut, les modèles de Kripke sont biens réflexifs, car chaque monde est en relation avec lui-même.

La symétrie intervient lorsque deux mondes sont en relation : si l'agent i ne peut pas différencier les mondes M et N depuis le monde M, il ne peut pas non plus différencier les mondes N et M depuis le monde N.

Pour la transitivité, elle sera rappelée plus tard dans le jeu des As et des huits, mais si un monde M est en relation avec un monde N, et que N est en relation avec un monde O, alors les mondes M et O sont également en relation.

Pour faire le parallèle avec les exemples précédents, la flèche rouge en figure 1 peut être notée comme « 1 R_A 2 », car Alice ne peut en effet pas différencier le mondes 1 du monde 2 si elle voit Bob avec de la boue sur le front.

On peut par ailleurs utiliser la notation v(φ, M), pour désigner la valeur de vérité de φ dans un monde M. Par exemple, v(As, 1) = 1, car Alice est sale dans le monde 1.
Mais v([A]As, 1) = 0, car dans le monde 1 Alice ne sait pas qu'elle est sale.

Pour résumer, on peut appliquer les connecteurs de la façon suivante : dans un monde donné ayant pour propriété φ, si un agent ne considère que ce monde à partir de celui-ci, alors on peut écrire [i]φ pour « l'agent i sait que φ ». Cependant, si dans ce même monde il considère plusieurs mondes où on a φ dans l'un et ¬φ dans l'autre, alors on écrira <i>φ, pour « l'agent i croit possible que φ ».

4 Mise en application avec des exemples

4.1 Enigme des 4 prisonniers et des chapeaux

L'énoncé est le suivant : quatre prisonniers condamnés à mort sont disposés en file indienne, ils regardent tous dans le même sens. Ils n'ont pas le droit de se retourner ou communiquer de quelque manière que ce soit. Un mur sépare le quatrième prisonnier des 3 autres. Le premier prisonnier peut alors voir le second et le troisième prisonnier, tandis que le second ne peut voir que le troisième, qui lui-même n'en voit aucun à cause du mur.

Chaque prisonnier a un chapeau sur la tête, qui peut être soit blanc soit noir. La seule information que possèdent les prisonniers à propos des chapeaux est le fait qu'ils savent qu'il y a deux chapeaux blancs et deux chapeaux noirs.
Leur bourreau décide de leur laisser une chance de survivre et leur propose un jeu : si l'un d'eux affirme la couleur de son chapeau, alors les prisonniers seront tous libres. S'il se trompe, ils seront exécutés. Il leur laisse une minute.

La question est la suivante : quel prisonnier annoncera la couleur de son chapeau ? [Figure 5].

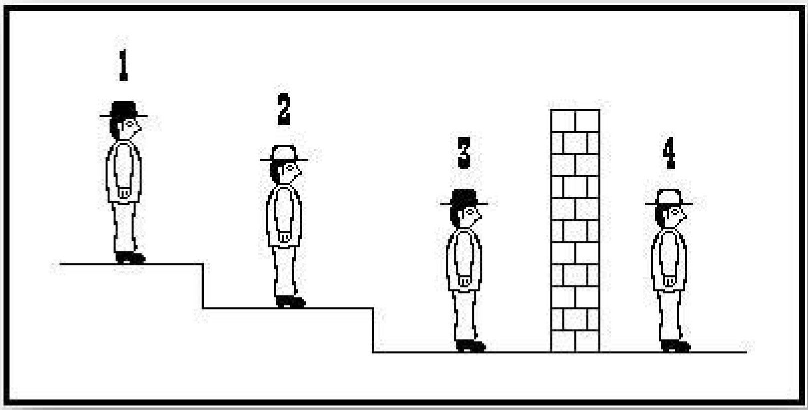

Figure 5 - Une configuration possible de l'énigme

Pour ce problème, on considèrera la notation « 1n » pour « le prisonnier 1 à un chapeau noir ». Pour un monde M donné, on notera alors « M ⊢ 1n, 2n » si les prisonniers 1 et 2 ont un chapeau noir dans le monde M, comme on n'a pas 3n et 4n, les prisonniers 3 et 4 ont donc un chapeau blanc. Sachant qu'il y a 2 chapeaux noirs parmi quatre hommes, il y a donc 6 permutations possibles des chapeaux. On sait donc qu'il existe 6 mondes possibles, qu'on notera avec des lettres pour éviter les confusions avec le numéro des prisonniers :

A ⊢ 1n, 2n
B ⊢ 1n, 3n
C ⊢ 1n, 4n
D ⊢ 2n, 3n
E ⊢ 2n, 4n
F ⊢ 3n, 4n ; soit le modèle de Kripke suivant [Figure 6, p11].

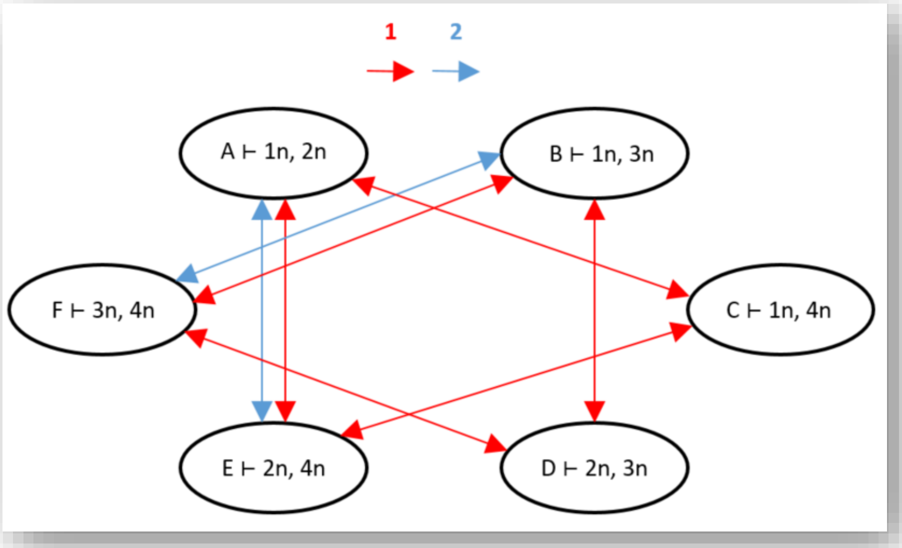

Figure 6 – Modèle de Kripke des 6 mondes possibles.

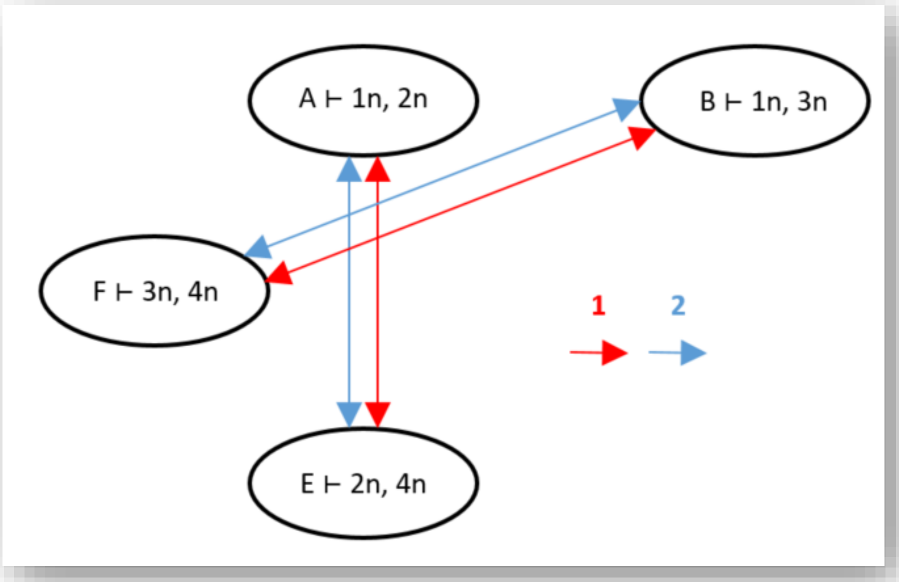

Figure 7 – Les 4 mondes possibles restant après que le premier prisonnier n'a pas parlé.

Dans la représentation ci-dessus, le prisonnier 1 est en bleu et le prisonnier 2 est en rouge. Les deux autres prisonniers sont omis car ils ne peuvent pas acquérir d'informations essentielles à la résolution du problème.

On constate que, dans un premier temps, le premier prisonnier est le seul à pouvoir déterminer son chapeau sans aucune information supplémentaire. En effet, si le prisonnier voit deux chapeau identiques, il se trouve alors dans un monde où il connait la couleur de son chapeau, car C ⊢ [1]1n (« Le prisonnier 1 sait que son chapeau est noir ») ou D ⊢ [1]¬1n (« Le prisonnier 1 sait que son chapeau n'est pas noir »), il peut donc directement annoncer la couleur de son chapeau.

On considère maintenant le cas où le premier prisonnier voit deux chapeaux de couleur différente. Contrairement au Muddy Children puzzle, il n'y a pas de « moment » où le bourreau pose sa question, ce qui annoncera aux autres qu'un prisonnier ne sait pas la couleur de son chapeau. En revanche, sachant que le temps imparti est d'une minute, on considère que si le premier prisonnier n'a pas parlé au bout d'une trentaine de secondes, alors il ne connait pas la couleur de son chapeau.

Passées ces 30 secondes, tous les prisonniers acquièrent alors la connaissance que le premier ne sait pas la couleur de son chapeau, à savoir ¬[1]1n. On a donc pour le prisonnier 2 : [2](¬[1]1n). Les mondes C et D peuvent être éliminés, car ces derniers forçaient [1]1n. Il reste alors quatre mondes possibles [Figure 7, p11].

Si on regarde les mondes restants, on pourrait croire que le fait que tout le monde sache que le premier prisonnier ne connaisse pas la couleur de son chapeau n'apporte rien, car le premier et le second prisonnier hésitent entre deux paires de mondes. Cependant, si l'on regarde de plus près, on peut voir que les deux paires de mondes sur lesquelles hésite le second prisonnier, soit A-E, et B-F, ont en commun le fait que le second prisonnier connaisse la couleur de son chapeau, car on a bien « (A ⊢ 2n) ∧ (E ⊢ 2n) » et « (F ⊢ ¬2n) ∧ (B ⊢ ¬2n) ». Or on sait également que le second prisonnier peut voir la couleur du chapeau du troisième prisonnier qui est devant lui. Si le troisième prisonnier a un chapeau noir, alors le second sait qu'il est dans le monde B ou F, donc que son propre chapeau est blanc et, raisonnement inverse, s'il voit que le chapeau du troisième prisonnier est blanc, ce qui implique que

- A ⊢ [2]2n
- E ⊢ [2]2n
- B ⊢ [2]¬2n
- F ⊢ [2]¬2n

Plus généralement, on peut donc considérer que [2](2n ∨ ¬2n) « le second prisonnier sait si son chapeau est noir ou blanc » à partir du moment où le premier prisonnier n'a pas parlé. Le second prisonnier doit donc simplement attendre quelque temps pour savoir si le premier prisonnier annonce ou non la couleur de son chapeau, pour pouvoir connaitre la couleur du sien.

4.2 Le jeu des As et des huits

Pour ce second exemple, nous allons voir le jeu des As et des huits [8]. Ce jeu se joue à trois joueurs avec 8 cartes : quatre As, et quatre huits. On va distribuer six des huit cartes en donnant deux cartes par joueur, il reste donc deux cartes qu'on laisse de côté, faces cachées. Les joueurs prennent alors leurs deux cartes et les mettent sur leur front, sans les regarder, de manière à ce que les deux autres joueurs puissent les voir. Le but du jeu est qu'un joueur devine les cartes qu'il a sur la tête pour gagner (l'ordre des cartes n'ayant pas d'importance). Chacun leur tour, les joueurs pourront dire s'ils savent ou non, quelles cartes ils ont au-dessus de leur tête. S'ils se trompent, ils perdent immédiatement la partie.

Nous allons au travers de ce jeu montrer comment la logique modale permet à un joueur de remporter la partie. Pour cela, nous allons analyser trois parties avec les joueurs suivants Arnaud, Bill et Claire.

Pour chaque partie, il y a un total de 19 mondes possibles, car 19 façons de repartir la « main » pour chaque joueur (paire d'As, paire de huits, ou As et huit) [Figure 8].

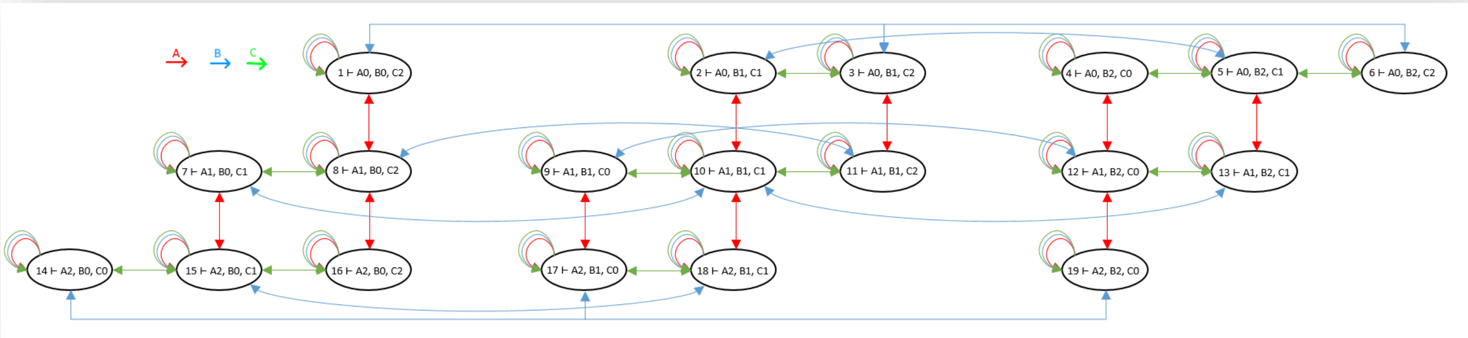

Figure 8 – Modèle de Kripke du jeu des As et des huits.

Ce modèle de Kripke semble au premier abord chaotique, mais nous allons l'expliquer pour fluidifier sa lecture. Les mondes sont numérotés de 1 à 19, la notation A1 correspond au joueur et au nombre d'As dans sa main ; 0 équivaut à une paire de huits, 1 à un As et un huit, et 2 pour une paire d'As.

La notation 1 ⊢ A0, B0, C2 signifie que le monde 1 force qu'Arnaud ait une paire de 8, Bill aussi, et Claire une paire d'As. Pour alléger le modèle en flèche, on considèrera qu'il est transitif, c'est-à-dire que si un joueur depuis le monde x considère le monde y, et que depuis le monde y il considère le monde z, alors le joueur considère le monde z depuis le monde x (sur le schéma, le monde 7 est alors considérable par Bill depuis le monde 13).

4.2.1 Première partie

Dans cette première partie, Arnaud commence. Il tient au-dessus de sa tête deux As, et dit qu'il ne connait pas ses cartes. Bill joue en second, il tient deux huits, et il ne peut lui non plus déterminer ses deux cartes. C'est au tour de Claire, nous allons voir comment cette dernière peut déterminer les cartes au-dessus de sa tête.

Après avoir vu les cartes d'Arnaud et de Bill, Claire considère trois mondes comme possibles : le 14, le 15 et le 16 [Figure 9]. En réalité, Arnaud considère possible les mondes 7 et 15, et Bill hésite entre les mondes 15 et 18, mais évidemment, Claire ne le sait pas. Après qu'Arnaud ait dit qu'il ne savait pas ses cartes, les deux joueurs obtiennent l'information que $\neg[A](A0 \lor A1 \lor A2)$, Bill et Claire peuvent donc éliminer le monde 14 des mondes possibles, car $14 \vdash [A]A2$.

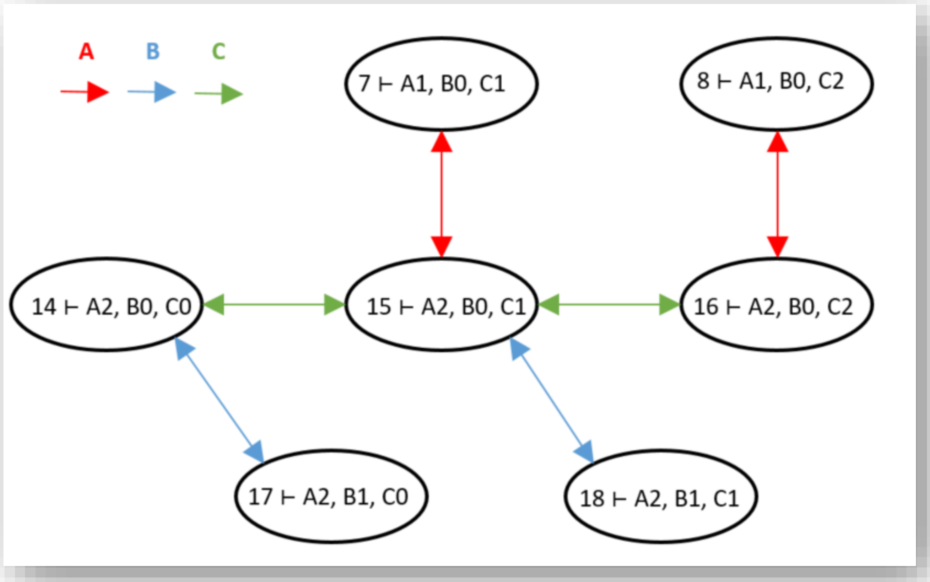

Figure 9 – Partie 1, mondes possibles du point de vue de Claire.

Cependant, cela n'aide pas Bill, car lors de son tour il n'est pas capable de déterminer ses cartes [Figure 10]. Avec cette information, Claire ne considère alors plus le monde 16 tel que $16 \vdash [B]B0$, ce qui laisse donc le monde $15 \vdash A2, B0, C1$, ainsi que $[C]C1$. Claire remporte donc cette première partie.

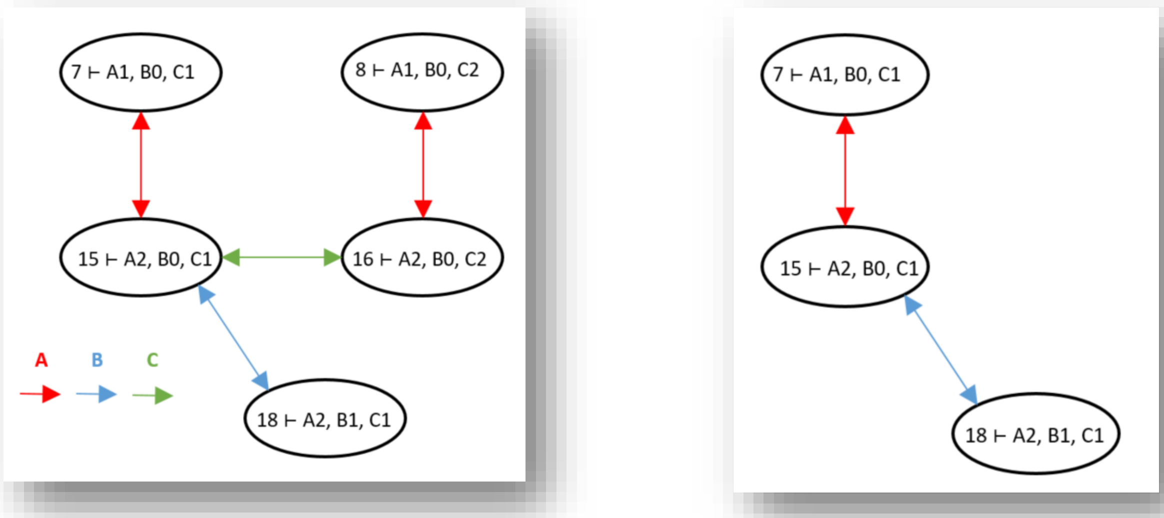

Figure 10 – Partie 1, avant et après le tour de Bill.

4.2.2 Seconde partie

Dans cette partie, nous suivons le point de vue d'Arnaud, qui ne peut pas déterminer ses cartes. Bill, qui tient deux huits, non plus. Claire n'est pas non plus apte à connaitre ses cartes, elle a un As et un huit. Arnaud peut à présent être sûr de sa main.

Reprenons depuis le début ; Arnaud connait les cartes de Bill et Claire (B2 C1), Il peut donc considérer uniquement le monde 5 et le 13 [Figure 11]. Lorsqu'il passe son tour, Bill et Claire peuvent éliminer les mondes où [A](A0 ∨ A1 ∨ A2), à savoir le monde 6 et le monde 14.

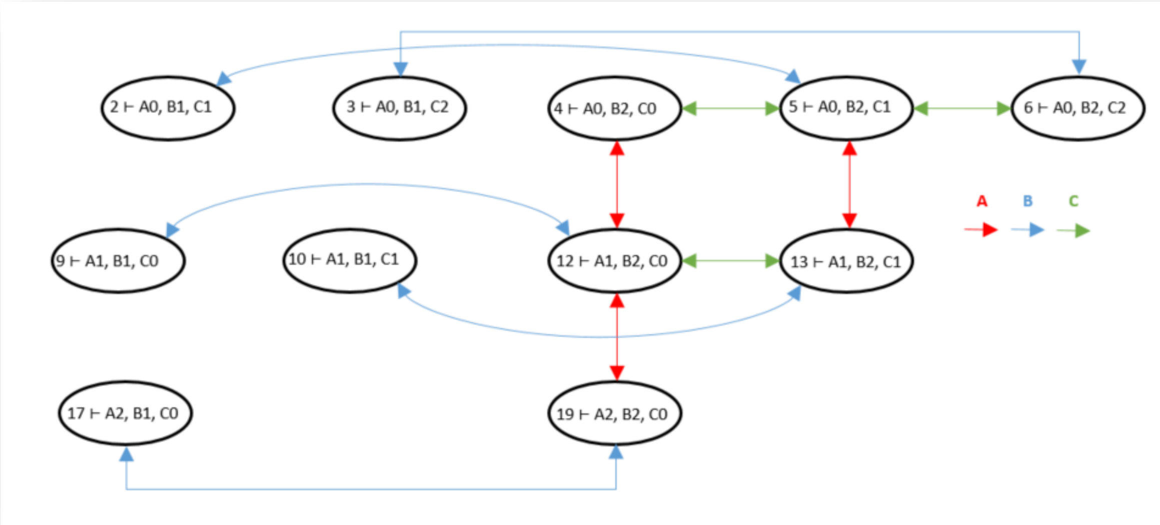

Figure 11 – Partie 2, modèle de Kripke simplifié et centré sur les mondes 5 et 13.

C'est ensuite Bill qui dit qu'il ne connait pas ses cartes, les mondes où Bill connait ses cartes ne sont alors plus envisagés par Arnaud et Claire, à savoir le monde 4 [Figure 12] (on rappelle que par transitivité, le monde 4 est en relation avec le monde 6 pour Claire).

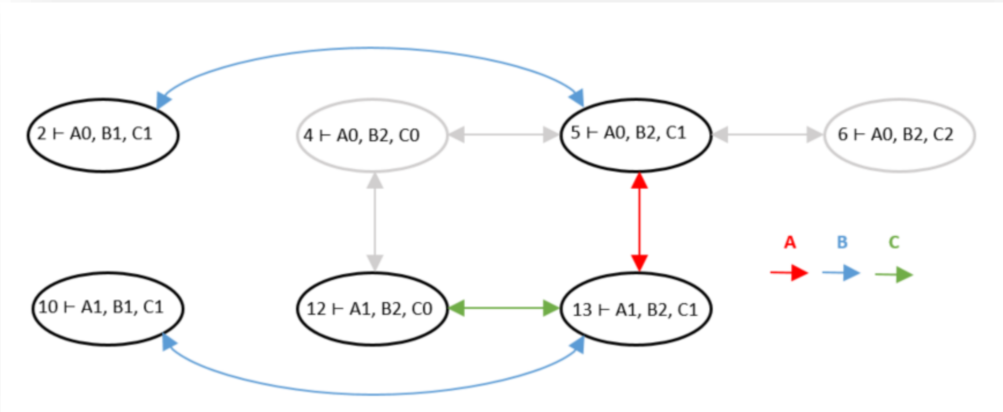

Figure 12 – Partie 2, modèle de Kripke simplifié après le tour de Bill.

C’est au tour de Claire. Elle non plus ne connait pas ses cartes. Arnaud peut alors éliminer les mondes où [C](C0 ∨ C1 ∨ C2), étant donné qu’elle ne considérait plus les mondes 4 et 6 avec les tours précédents, on peut enlever le monde 5 des mondes considérés par Arnaud.

C’est donc au tour d’Arnaud. Comme il hésitait entre les mondes 5 et 13, et que le monde 5 vient d’être considéré comme non possible, alors il ne lui reste que le monde 13 de possible. Arnaud remporte la partie en annonçant qu’il a dans sa main un As et un huit.

4.2.3 Troisième partie

Pour la dernière partie, nous suivons Bill. Arnaud et Claire tiennent tous deux un As et un huit. Personne ne peut déduire ses cartes au premier tour.

Au début du second tour, Arnaud dit qu’il ne sait pas.

Pour le premier tour, nous allons successivement enlever les mondes 14 et 6 après Arnaud, 16 et 4 après Bill, et les mondes 1, 5, 15 et 19 après que Claire ait parlé [Figure 13].

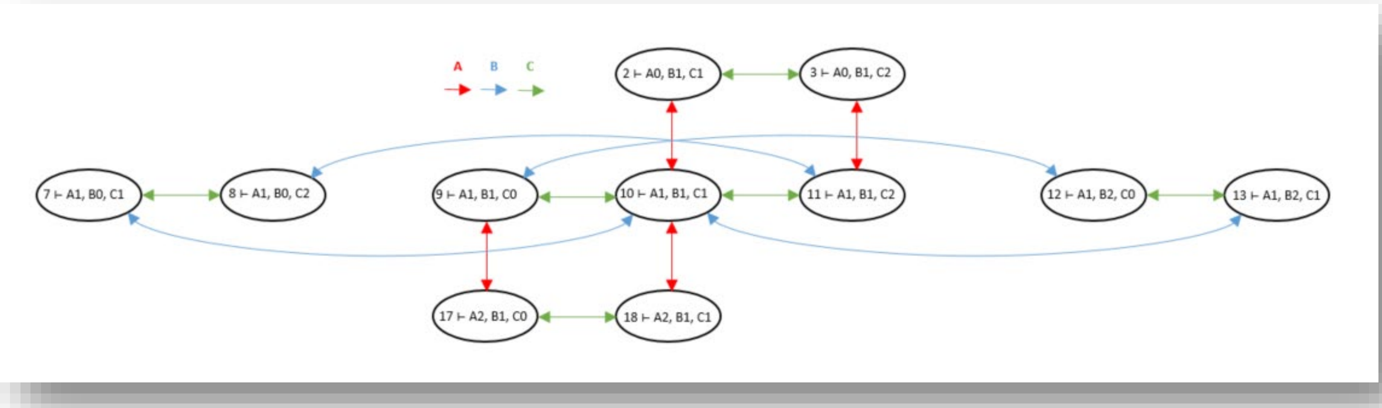

Figure 13 – Partie 3, modèle de Kripke après le premier tour

C’est à nouveau au tour d’Arnaud de parler. Lui et Bill ont en commun le fait qu’ils connaissent tous deux le jeu de Claire (C1), ce qu’on peut d’ailleurs représenter par [A][B]C1, ce qui correspond aux mondes 2, 7, 10, 13 et 18 [Figure 14].

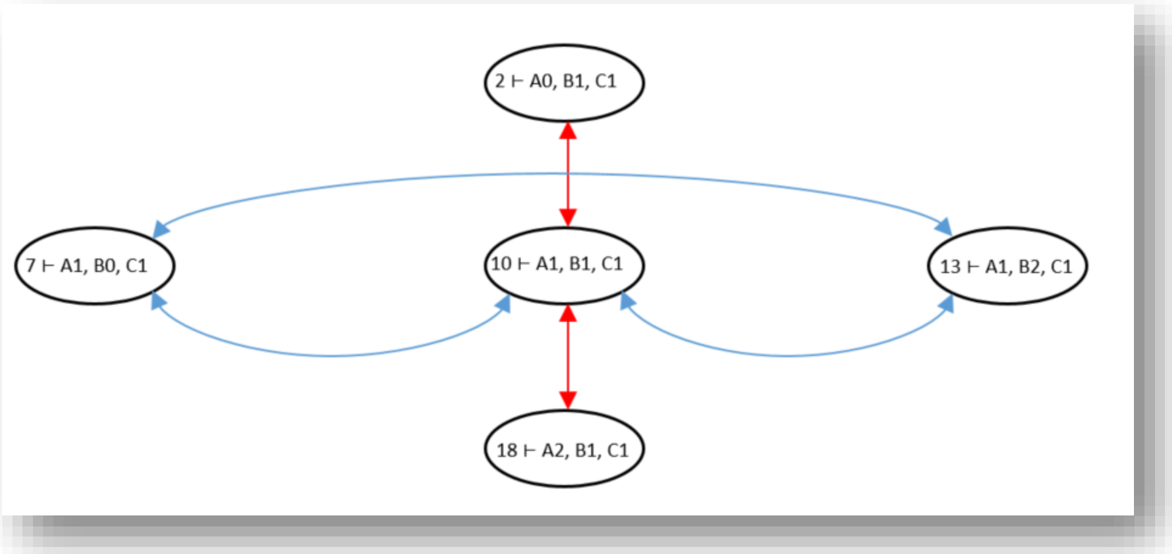

Figure 14 – Partie 3, modèle de Kripke simplifié entre Arnaud et Bill.

Mais Arnaud connait les cartes de Bill, il hésite donc seulement entre les mondes 2, 10 et 18, mais cela ne lui permet pas de savoir ses cartes, il dit alors qu'il ne sait pas. C'est le second tour de Bill, et il vient d'apprendre qu'Arnaud ne connait pas ses cartes, donc qu'il ne peut pas se trouver dans les mondes 7 et 13. Comme Bill connait les cartes d'Arnaud, il conclut que le seul monde possible restant est le monde 10. Bill remporte la dernière partie en annonçant qu'il a en main un As et un huit.

Nous avons donc pu voir au travers de ces trois parties que la logique modale nous permettait de comprendre quel joueur pouvait déduire ses cartes au fur et à mesure de l'avancée de la partie, notamment à l'aide des modèles de Kripke qui améliorent la compréhension, car étant plus visuels et donc plus intuitifs.

5 Bilan du stage

5.1 Résultats obtenus et problèmes rencontrés

Le but de ce stage était dans un premier temps de comprendre la modélisation et le traitement du Muddy Children Puzzle en logique épistémique, et de comprendre la sémantique des mondes possibles. Dans un second temps, trouver de nouveaux exemples, pour pouvoir appliquer les connaissances apprises en essayant d'expliquer le mieux possible à quelqu'un qui ne serait pas initié à cette logique. A la fin de ce stage, j'ai acquis une bonne compréhension des mondes de Kripke, de ce type de problème, et plus généralement j'estime être en capacité de présenter les bases de la logique épistémique à une personne non familière avec ce type de logique.

Nous devions cependant être trois à faire ce stage, mais malheureusement mes deux collègues n'ont pas pu y participer. Cela avait son lot de points positifs et négatifs ; dans un premier temps être en groupe pour apprendre peut s'avérer synergique, étant donné que certains points clés qu'un membre ne comprend pas peuvent être expliqués par un autre membre qui a compris, c'est en ce sens qu'être seul a constitué un désavantage dans l'apprentissage de la logique modale. Cependant, cela m'a aussi permis d'avoir des rendez-vous plus fréquents, car il n'y avait pas besoin que nous soyons tous d'accord et disponibles pour les dates où nous devions nous voir, ce qui a facilité nos rencontres et nos échanges avec M. Rétoré et M. Catta (Doctorant LIRMM). Un autre avantage à être seul est le fait qu'il n'y a pas besoin de se mettre d'accord pour des décisions, ce qui constitue un gain de temps.

5.2 Bilan personnel et perspectives d'avenir

Je souhaiterais conclure en disant que ce stage m'a été fortement bénéfique dans plusieurs domaines. Tout d'abord j'ai découvert les bases de la logique modale ainsi que de la logique épistémique, ce qui m'a fait découvrir une branche très intéressante des mathématiques qui m'a beaucoup intéressé ; mais cela m'a permis également d'élargir mes recherches pour commencer à me renseigner sur la logique du premier ordre, qui sera au programme de la troisième année de licence.

Pour ce qui est de mon projet professionnel, ma vision de l'avenir reste relativement floue, et je n'ai pas encore une idée précise du MASTER informatique que je souhaite faire, mais ce stage et la découverte de logiques autre que classique, avec la part de mathématiques qui s'y trouvent me font pencher vers un master plus orienté mathématiques/algorithmique.

6 Bibliographie

[1] Article Logique Modale de Wikipédia wikipedia.org/Logique_Modale

[2] L'exemple des muddy children est tiré de "Reasoning about knowledge" [3], lui-même adapté du problème des « unfaithtful wifes » (femmes infidèles) du livre "A mathematician's miscellany" [4] de John Edensor Littlewood.

[3] “Reasoning about Knowledge” par R. Fagin, J.Y. Halpern, Y. Moses, M.Y. Vardi (1995), chapitre 2.

[4] "A mathematician's miscellany" par John Edensor Littlewood (1953).

[5] “Dynamic Epistemic Logic”, Stanford Encyclopedia of Philosophy Archive, par Baltag, Alexandru et Bryan Renne, 2016 : plato.stanford.edu/dynamic-epistemic/

[6] “Modal Logic”, Stanford Encyclopedia of Philosophy Archive, par Garson, James, 2000 : plato.stanford.edu/logic-modal

[7] “KRIPKE SAUL”, Encyclopædia Universalis, Scott SOAMES : universalis.fr/saul-kripke

[8] Chapitre 1 de “Dynamic Epistemic Logic” [5], Introduction and Overview, page 11.

[9] “La logique pas à pas”, Jacques Duparc, 2015. Intro de la partie II : « En route vers les mondes possibles »

Logos et images

Logo Université Montpellier : umontpellier.fr

Logo Faculté des Sciences : sciences.edu.umontpellier.fr

Logo Réseau Figure : reseau-figure.fr

Logo LIRMM : lirmm.fr

Image des 4 prisonniers : blogspot.com