Intelligence in Strategic Games (Extended Abstract)*

Pavel Naumov[1] , **Yuan Yuan**[2]
[1]University of Southampton, the United Kingdom
[2]Vassar College, the United States
p.naumov@soton.ac.uk, yyuan@vassar.edu

Abstract

If an agent, or a coalition of agents, has a strategy, knows that she has a strategy, and knows what the strategy is, then she has a know-how strategy. Several modal logics of coalition power for know-how strategies have been studied before.

The contribution of the article is three-fold. First, it proposes a new class of know-how strategies that depend on the intelligence information about the opponents' actions. Second, it shows that the coalition power modality for the proposed new class of strategies cannot be expressed through the standard know-how modality. Third, it gives a sound and complete logical system that describes the interplay between the coalition power modality with intelligence and the distributed knowledge modality in games with imperfect information.

1 Introduction

The Battle of the Atlantic was a classical example of the matching pennies game. British (and American) admirals were choosing routes of the allied convoys and the Germans picked routes of their U-boats. If their trajectories crossed, the Germans scored a win, if not the allies did. Neither of the players appeared to have a *strategy* that would guarantee a victory.

The truth, however, was that during the most of the battle one of the sides had exactly such a strategy. First, it was the British who broke German Enigma cipher in summer of 1941. Although the Germans did not know about the British success, they changed codebook and added fourth wheel to Enigma in February 1942 thus preventing the British from decoding German messages. The very next month, in March 1942, German navy cryptography unit, B-Dienst, broke the allied code and got access to convoy route information. The Germans lost their ability to read allied communication in December 1942 due to a routine change in the allied codebook. The same month, the British were able to read German communication as a result of capturing codebook from a U-boat in Mediterranean. In March 1943, the Germans changed codebook again and, unknowingly, disabled British ability to read German messages. Simultaneous, the Germans caught up and started to decipher British transmissions again [Budiansky, 2002; Showell, 2003].

At almost any moment during these two years one of the sides was able to read the communications of the other side. However, the two sides never have been able to read each other's messages at the same time to notice that the other side knows more than it should have known. As a result, neither of them was able to figure out that their own code is insecure. Finally, in May 1943, with the help of US Navy, the British cracked German messages while the Germans still were reading British. It was the first time the allies understood that their code was insecure. A new convoy cipher was immediately introduced and the Germans have never been able to break it again, while the allies continued reading Enigma-encrypted transmissions till the end of the war [Budiansky, 2002].

In this article, we study coalition power in strategic games assuming that the coalition has intelligence information about the moves of all or some of its opponents. Throughout the article, we refer to the information about the future actions of the other agents simply as "the intelligence". We write $[C]_I\varphi$ if a coalition C has a strategy to achieve an outcome φ in one step as long as the coalition has the intelligence about the move of each agent in set I. For example,

$$[British]_{Germans}(\text{Convoy is saved}).$$

Strategic games could be generalized to strategic games with imperfect information. Unlike perfect information strategic games, the initial state in an imperfect information game could be unknown to the players. For example, consider a hypothetical setting in which a British convoy and a German U-boat have to choose between three routes from point A to point B: route 1, route 2, and route 3, see Figure 1. Let us furthermore assume that it is known to both sides that one of these routes is blocked by Russian naval mines. Although the mines are located along route 1, neither the British nor the Germans know this. If the British have an access to the intelligence about German U-boats, then, in theory, they have a strategy to save the convoy. For example, if the Germans will use route 2, then the British can use route 3. However, since the British do not know the location of the Russian mines, even after they receive information about German plans, they still would not know how to save the convoy.

*This paper is an extended abstract of an article in Journal of Artificial Intelligence Research [Naumov and Yuan, 2021].

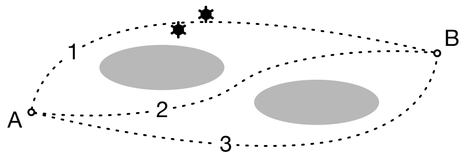

Figure 1: Three routes from point A to point B.

Modality $[C]_{\varnothing}\varphi$ is the coalitional power modality proposed by Marc Pauly [Pauly, 2001; Pauly, 2002]. He gave a sound and complete axiomatization of this modality in the case of perfect information strategic games. It has been suggested in several recent works that, in the case of the games with imperfect information, strategic power modality in Marc Pauly logic should be restricted to existence of *know-how*[1] strategies [Ågotnes and Alechina, 2019; Naumov and Tao, 2017a; Fervari *et al.*, 2017; Naumov and Tao, 2018b; Naumov and Tao, 2018c; Naumov and Tao, 2018a; Cao and Naumov, 2020]. That is, modality $[C]\varphi$ should stand for "coalition C has a strategy, it distributively knows that it has a strategy, and it distributively knows what the strategy is". In this article we adopt this approach to strategic power with intelligence. For example, in the imperfect information setting depicted in Figure 1, after receiving the intelligence report, the British have a strategy, they know that they have a strategy, but they *do not know* what the strategy is:

$$\neg[British]_{Germans}(\text{Convoy is saved}).$$

At the same time, since the Russians presumably know the location of their mines,

$$[\text{British, Russians}]_{\text{Germans}}(\text{Convoy is saved}).$$

This article contains two main technical results. First, we show that know-how with intelligence modality $[C]_I\varphi$ cannot be defined through the standard know-how modality. Second, we give a complete logical system that describes the interplay between the coalition power with intelligence modality $[C]_I$ and the distributed knowledge modality K_C in the imperfect information setting. The most interesting axiom of our system is a generalized version of Marc Pauly's [Pauly, 2001; Pauly, 2002] Cooperation axiom that connects intelligence I and coalition C parameters of the modality $[C]_I$.

2 Games with Imperfect Information

For any sets X and Y, by X^Y we mean the set of all functions from Y to X. Throughout the article we assume a fixed (possibly infinite) set of agents $\mathcal{A}$ and a nonempty set of propositional variables.

Definition 1. *A game is a tuple* $(W, \{\sim_a\}_{a\in\mathcal{A}}, \Delta, M, \pi)$, *where*

1. *W is a set of states,*
2. *$\sim_a$ is an "indistinguishability" equivalence relation on set W for each agent $a \in \mathcal{A}$,*
3. *Δ is a nonempty set, called the "domain of actions",*
4. *$M \subseteq W \times \Delta^{\mathcal{A}} \times W$ is a relation called "aggregation mechanism",*
5. *π is a function that maps propositional variables to subsets of W.*

A function δ from set $\Delta^{\mathcal{A}}$ is called a *complete action profile*.

Figure 2 depicts a diagram of the Battle of the Atlantic game with imperfect information as described in the introduction. For the sake of simplicity, we treat the British, the Germans, and the Russians as single agents, not groups of agents. The game has five states: 1, 2, 3, s, and d. States 1, 2, and 3 are three "initial" states that correspond to possible locations of Russian mines along route 1, route 2, or route 3. Neither the British nor the Germans can distinguish these states, which is shown in the diagram by labels on the dashed lines connecting these three states. The Russians know location of the mines, and, thus, they can distinguish these states. The other two states are "final" states s and d that describe if the convoy made it safe (s) or was destroyed (d) by either a U-boat or a mine. The designation of some states as "initial" and others as "final" is specific to the Battle of the Atlantic game. In general, our Definition 1 does not distinguish between such states and we allow games to take multiple consecutive transitions from one state to another.

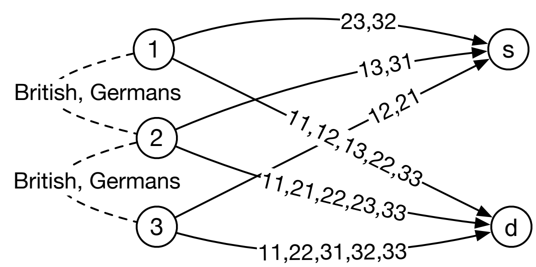

Figure 2: Battle of the Atlantic with imperfect information.

The domain of actions Δ in this game is $\{1, 2, 3\}$. For the British and the Germans actions represent the choice of routes that they make for their convoys and U-boats respectively. The Russians are passive players in this game. Their action does not affect the outcome of the game. Technically, a complete action profile is a function δ from set $\{\text{British}, \text{Germans}, \text{Russians}\}$ into set $\{1, 2, 3\}$. Since, there are only three players in the Battle of the Atlantic game, it is more convenient to represent function δ by triple $bgr \in \{1, 2, 3\}^3$, where b is the action of the British, g is the action of the Germans, and c is the action of the Russians.

The mechanism M of the Battle of the Atlantic game is captured by the directed edges in Figure 2 labeled by complete action profiles. Since value r in a profile bgr does not affect the outcome, it is omitted on the diagram. For example, directed edge from state 1 to state s is labeled with 23

[1]Know-how strategies were studied before under different names. While Jamroga and Ågotnes talked about "knowledge to identify and execute a strategy" 2007, Jamroga and van der Hoek discussed "difference between an agent knowing that he has a suitable strategy and knowing the strategy itself" 2004. Van Benthem called such strategies "uniform" 2001. Wang gave a complete axiomatization of "knowing how" as a binary modality 2015; 2018, but his logical system does not include the knowledge modality.

and 32. This means that the mechanism M contains triples $(1, 231, s)$, $(1, 232, s)$, $(1, 233, s)$, $(1, 321, s)$, $(1, 322, s)$, and $(1, 323, s)$.

The definition of a game that we use here is more general than the one used in the original Marc Pauly's semantics of the logic of coalition power. Namely, we assume that the mechanism is a relation, not a function. On one hand, this allows us to talk about nondeterministic games where for each initial state and each complete action profile there might be more than one outcome. On the other hand, this also allows no outcome to exist for some combinations of the initial states and the complete action profiles. If in a given state under a given complete action profile there is no next state, then we interpret this as termination of the game. This resembles a user choosing to quit a computer application: once the quit button is pressed, the application terminates without reaching any new state. If needed, games with termination can be excluded and an additional axiom $\neg[C]_{\varnothing}\bot$ be added to the logical system. The proof of the completeness will remain mostly unchanged.

Definition 2. *For any states $w, w' \in W$ and any coalition C, let $w \sim_C w'$ if $w \sim_a w'$ for each agent $a \in C$.*

In particular, $w \sim_{\varnothing} w'$ for any two states of the game.

Lemma 1. *For any coalition C, relation $\sim_C$ is an equivalence relation on set W.* □

3 Syntax and Semantics

In this section, we define the syntax and the semantics of our formal system. By a coalition we mean any finite subset of the set of all agents $\mathcal{A}$. Finiteness of coalitions will be important for the proof of the completeness.

Definition 3. *Let Φ be the minimal set of formulae such that*

1. *$p \in \Phi$ for each propositional variable p,*
2. *$\varphi \to \psi, \neg\varphi \in \Phi$ for all $\varphi, \psi \in \Phi$,*
3. *$\mathsf{K}_C\varphi \in \Phi$ for each formula $\varphi \in \Phi$ and each coalition $C \subseteq \mathcal{A}$,*
4. *$[C]_B\varphi \in \Phi$ for each formula $\varphi \in \Phi$ and all **disjoint** coalitions B, C.*

In other words, the language of our logical system is defined by grammar:

$$\varphi := p \mid \neg\varphi \mid \varphi \to \varphi \mid \mathsf{K}_C\varphi \mid [C]_B\varphi,$$

where coalitions C and B are disjoint. Formula $\mathsf{K}_C\varphi$ stands for "coalition C distributively knows φ" and formula $[C]_B\varphi$ for "coalition C distributively knows strategy to achieve φ as long as it gets the intelligence on actions of coalition B". We assume that Boolean constants $\top$ and $\bot$ are defined through negation, implication, and a propositional variable in the standard way.

By an action profile of a coalition C we mean any function from set Δ^C. For any two functions f, g, we write $f =_X g$ if $f(x) = g(x)$ for each $x \in X$.

Next is the key definition of this article. Its part 5 gives the semantics of modality $[C]_B$. This part uses state w' to capture the fact that the strategy succeeds in each state indistinguishable by coalition C from the current state w. In other words, the coalition C *knows* that this strategy will succeed. Except for the addition of coalition B and its action profile β, this is essentially the same definition as the one used in [Ågotnes and Alechina, 2019; Naumov and Tao, 2017a; Fervari *et al.*, 2017; Naumov and Tao, 2017b; Naumov and Tao, 2018c; Naumov and Tao, 2018b; Naumov and Tao, 2018a].

Definition 4. *For any game $(W, \{\sim_a\}_{a\in\mathcal{A}}, \Delta, M, \pi)$, any state $w \in W$, and any formula $\varphi \in \Phi$, let satisfiability relation $w \Vdash \varphi$ be defined as follows:*

1. *$w \Vdash p$, if $w \in \pi(p)$, where p is a propositional variable,*
2. *$w \Vdash \neg\varphi$, if $w \nVdash \varphi$,*
3. *$w \Vdash \varphi \to \psi$, if $w \nVdash \varphi$ or $w \Vdash \psi$,*
4. *$w \Vdash \mathsf{K}_C\varphi$, if $w' \Vdash \varphi$ for each $w' \in W$ such that $w \sim_C w'$,*
5. *$w \Vdash [C]_B\varphi$, if for any action profile $\beta \in \Delta^B$ of coalition B there is an action profile $\gamma \in \Delta^C$ of coalition C such that for any complete action profile $\delta \in \Delta^{\mathcal{A}}$ and any states $w', u \in W$ if $\beta =_B \delta$, $\gamma =_C \delta$, $w \sim_C w'$, and $(w', \delta, u) \in M$, then $u \Vdash \varphi$.*

For example, for the game depicted in Figure 2,

$$1 \Vdash [\text{British, Russians}]_{\text{Germans}}(\text{Convoy is saved}).$$

Indeed, statement $1 \sim_{\text{British, Russians}} w'$ is true only for one state $w' \in W$, namely state 1 itself. Then, for any action profile $\beta \in \{1, 2, 3\}^{\{\text{Germans}\}}$ of the single-member coalition {Germans} we can define action profile $\gamma \in \{1, 2, 3\}^{\{\text{British, Russians}\}}$ as, for example,

$$\gamma(a) = \begin{cases} 3, & \text{if } a = \text{British and } \beta(\text{Germans}) = 2, \\ 2, & \text{if } a = \text{British and } \beta(\text{Germans}) = 3, \\ 1, & \text{if } a = \text{Russians.} \end{cases}$$

In other words, if profile β assigns the Germans route 2, then profile γ assigns the British route 3 and vice versa. Assignment of an action to the Russians in not important. This way, no matter what the Germans' action is, the British convoy will avoid both the German U-boat and the Russian mines in the game that starts from state $w' = 1$. At the same time,

$$1 \Vdash \neg[\text{British}]_{\text{Germans}}(\text{Convoy is saved}).$$

because the British cannot distinguish states 1, 2, and 3 without the Russians. In other words, $1 \sim_{\text{British}} w'$ for any state $w' \in \{1, 2, 3\}$. Thus, for each action profile $\beta \in \{1, 2, 3\}^{\{\text{Germans}\}}$ we need to have a single action profile $\gamma \in \{1, 2, 3\}^{\{\text{British, Russians}\}}$ that would bring the convoy to state s from any of the states 1, 2, and 3. Such profile γ does not exist because, even if the British know where the Germans U-boat will be, there is no single strategy to choose path that would avoid Russian mines from all three indistinguishable states 1, 2, and 3.

Note that item 4 of Definition 4 interprets the knowledge modality K_C as *distributed* knowledge of coalition C. Similarly, the knowledge of "how", implicitly referred to by item 5 of the same definition through indistinguishability relation $\sim_C$, is also distributed knowledge of coalition C. Thus, even if a coalition knows how to achieve φ, it is possible that no single individual member of the coalition might know how to do this. In order for the coalition to execute a strategy that it knows, its members might need to communicate with each other to "assemble" the distributed knowledge. The mechanism and the result of such a communication is not a trivial issue. For example, if a statement φ is distributively known to a coalition, then φ might even be false after such a communication [Ågotnes and Wáng, 2017]. In this article, following the existing tradition in the logics of distributively knowing strategy, we investigate the properties of distributed knowledge on an abstract level, without considering the communication between the agents required to execute the strategy. Of course, one can potentially consider the other forms of group knowledge such as individual knowledge (each agent in the group knows) and common knowledge. However, know-how strategies based on either of the last two forms of knowledge do not satisfy the Cooperation axiom listed in Section 4. As a result, they are significantly less interesting from the logical point of view.

Finally, observe that the formula $\mathsf{K}_\varnothing\varphi$ means that statement φ is satisfied in all states of the game. Modality $\mathsf{K}_\varnothing$ is sometimes referred to as universal modality.

4 Axioms

In this section, we present our formal logical system, Know-How Logic with the Intelligence, for reasoning about the interplay between distributed knowledge modality K_C and know-how with intelligence modality $[C]_B$.

Definition 5. *In addition to the propositional tautologies in language Φ, the axioms of the Know-How Logic with the Intelligence are*

1. *Truth:* $\mathsf{K}_C\varphi \to \varphi$,
2. *Distributivity:* $\mathsf{K}_C(\varphi \to \psi) \to (\mathsf{K}_C\varphi \to \mathsf{K}_C\psi)$,
3. *Negative Introspection:* $\neg\mathsf{K}_C\varphi \to \mathsf{K}_C\neg\mathsf{K}_C\varphi$,
4. *Epistemic Monotonicity:* $\mathsf{K}_C\varphi \to \mathsf{K}_D\varphi$, *where* $C \subseteq D$,
5. *Strategic Introspection:* $[C]_B\varphi \to \mathsf{K}_C[C]_B\varphi$,
6. *Empty Coalition:* $\mathsf{K}_\varnothing\varphi \to [\varnothing]_\varnothing\varphi$,
7. *Cooperation: for any pairwise disjoint sets* B, C, *and* D, $[C]_B(\varphi \to \psi) \to ([D]_{B,C}\varphi \to [C,D]_B\psi)$,
8. *Intelligence Monotonicity:* $[C]_B\varphi \to [C]_{B'}\varphi$, *where* $B \subseteq B'$,
9. *None to Analyze:* $[\varnothing]_B\varphi \to [\varnothing]_\varnothing\varphi$.

Note that in the Cooperation axiom above and often throughout the rest of the article we abbreviate $B\cup C$ as B, C. However, we keep writing $B \cup C$ when notation B, C could be confusing.

Next, we define two derivability relations. A unary relation $\vdash \varphi$ and a binary relation $X \vdash \varphi$, where $X \subseteq \Phi$ is a set of formulae and $\varphi \in \Phi$ is a formula.

Definition 6. $\vdash \varphi$ *if there is a finite sequence of formulae that ends with φ such that each formula in the sequence is either an axiom or could be obtained from the previous formulae using the Modus Ponens, the Epistemic Necessitation, or the Strategic Necessitation inference rules:*

$$\frac{\varphi, \quad \varphi \to \psi}{\psi}, \qquad \frac{\varphi}{\mathsf{K}_C\varphi}, \qquad \frac{\varphi}{[C]_B\varphi}.$$

If $\vdash \varphi$, then we say that φ is a *theorem* of our logical system.

Definition 7. $X \vdash \varphi$ *if there is a finite sequence of formulae that ends with φ such that each formula in the sequence is either a **theorem**, or an element of X, or could be obtained from the previous formulae using **only** the Modus Ponens inference rules.*

Note that if set X is empty, then statement $X \vdash \varphi$ is equivalent to $\vdash \varphi$. Thus, instead of $\varnothing \vdash \varphi$ we write $\vdash \varphi$. We say that set X is consistent if $X \nvdash \bot$.

5 Main Results

By Φ^- we denote the fragment of language Φ in which modality in occurrences of modality $[B]_C\varphi$ set C is empty. The proofs of the following results appear in the full version of this paper [Naumov and Yuan, 2021].

Theorem 1 (undefinability)**.** *The modality $[C]_B\varphi$ is not definable in language Φ^-.*

Theorem 2 (soundness)**.** *For any state $w \in W$ of any game $(W, \{\sim_a\}_{a\in\mathcal{A}}, \Delta, M, \pi)$, any set of formulae $X \subseteq \Phi$, and any formula $\varphi \in \Phi$, if $w \Vdash \chi$ for each formula $\chi \in X$ and $X \vdash \varphi$, then $w \Vdash \varphi$.*

Theorem 3 (completeness)**.** *If $Y \nvdash \varphi$, then there is a state w of a game such that $w \Vdash \chi$ for each $\chi \in Y$ and $w \nVdash \varphi$.*

6 Conclusion

In this article, we proposed the notion of a strategy with intelligence, proved that corresponding know-how modality is not definable through the standard know-how modality, and gave a sound and complete axiomatization of a bimodal logic that describes the interplay between the strategic power with intelligence and the distributed knowledge modalities in the setting of strategic games with imperfect information.

A natural question is decidability of the proposed logical system. Unfortunately, the standard filtration technique [Gabbay, 1972] can not be easily applied here to produce a finite model. Indeed, it is crucial for the proof of the completeness that for each action there is another action with a higher value of the "key" component. Thus, for the proposed construction to work, the domain of choices must be infinite. One perhaps might be able to overcome this by changing the second component of the action from an infinite linear ordered set to a finite circularly "ordered" set as in rock-paper-scissors game, but we have not been able to prove this.

References

[Ågotnes and Alechina, 2019] Thomas Ågotnes and Natasha Alechina. Coalition logic with individual, distributed and common knowledge. *Journal of Logic and Computation*, 29:1041–1069, 11 2019.

[Ågotnes and Wáng, 2017] Thomas Ågotnes and Yì N Wáng. Resolving distributed knowledge. *Artificial Intelligence*, 252:1–21, 2017.

[Budiansky, 2002] Stephen Budiansky. German vs. allied codebreakers in the battle of the atlantic. *International Journal of Naval History*, 1(1), April 2002.

[Cao and Naumov, 2020] Rui Cao and Pavel Naumov. Knowing the price of success. *Artificial Intelligence*, 284:103287, 2020.

[Fervari *et al.*, 2017] Raul Fervari, Andreas Herzig, Yanjun Li, and Yanjing Wang. Strategically knowing how. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pages 1031–1038, 2017.

[Gabbay, 1972] Dov M Gabbay. A general filtration method for modal logics. *Journal of Philosophical Logic*, 1(1):29–34, 1972.

[Jamroga and Ågotnes, 2007] Wojciech Jamroga and Thomas Ågotnes. Constructive knowledge: what agents can achieve under imperfect information. *Journal of Applied Non-Classical Logics*, 17(4):423–475, 2007.

[Jamroga and van der Hoek, 2004] Wojciech Jamroga and Wiebe van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, 63(2-3):185–219, 2004.

[Naumov and Tao, 2017a] Pavel Naumov and Jia Tao. Coalition power in epistemic transition systems. In *Proceedings of the 2017 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 723–731, 2017.

[Naumov and Tao, 2017b] Pavel Naumov and Jia Tao. Together we know how to achieve: An epistemic logic of know-how. In *Proceedings of 16th conference on Theoretical Aspects of Rationality and Knowledge (TARK), July 24-26, 2017, EPTCS 251*, pages 441–453, 2017.

[Naumov and Tao, 2018a] Pavel Naumov and Jia Tao. Second-order know-how strategies. In *Proceedings of the 2018 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 390–398, 2018.

[Naumov and Tao, 2018b] Pavel Naumov and Jia Tao. Strategic coalitions with perfect recall. In *Proceedings of Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[Naumov and Tao, 2018c] Pavel Naumov and Jia Tao. Together we know how to achieve: An epistemic logic of know-how. *Artificial Intelligence*, 262:279 – 300, 2018.

[Naumov and Yuan, 2021] Pavel Naumov and Yuan Yuan. Intelligence in strategic games. *Journal of Artificial Intelligence Research*, 71:521–556, 2021.

[Pauly, 2001] Marc Pauly. *Logic for Social Software*. PhD thesis, Institute for Logic, Language, and Computation, 2001.

[Pauly, 2002] Marc Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.

[Showell, 2003] Jak P. Mallmann Showell. *German Naval Code Breakers*. Naval Inst Pr, 1st edition, 2003.

[van Benthem, 2001] Johan van Benthem. Games in dynamic-epistemic logic. *Bulletin of Economic Research*, 53(4):219–248, 2001.

[Wang, 2015] Yanjing Wang. A logic of knowing how. In *Logic, Rationality, and Interaction*, pages 392–405. Springer, 2015.

[Wang, 2018] Yanjing Wang. A logic of goal-directed knowing how. *Synthese*, 195(10):4419–4439, 2018.