Available online at www.sciencedirect.com

SCIENCE DIRECT®

Theoretical Computer Science 343 (2005) 97–113

Theoretical Computer Science

www.elsevier.com/locate/tcs

Non-cooperative computation: Boolean functions with correctness and exclusivity

Yoav Shoham[a,*], Moshe Tennenholtz[b]

[a] *Computer Science Department, Stanford University, Stanford, CA 94305, USA*
[b] *Faculty of Industrial Engineering and Management, Technion–Israel Institute of Technology, Haifa 32000, Israel*

Abstract

We introduce the concept of *non-cooperative computation* (*NCC*), which is the joint computation of a function by self-motivated agents, where each of the agents possesses one of the inputs to the function. In NCC the agents communicate their input (truthfully or not) to a trusted center, which performs a commonly-known computation and distributes the results to the agents. The question is whether the agents can be incented to communicate their true input to the center, allowing all agents to compute the function correctly. NCC is a game theoretic concept and specifically is couched in terms of mechanism design. NCC is a very broad framework and is specialized by imposing specific structure on the agents' utility functions. The technical results we present are specific to the setting in which each agent has a primary interest in computing the function and a secondary interest in preventing the others from computing it (properties called *correctness* and *exclusivity*). For this setting we provide a complete characterization of the Boolean functions that are non-cooperatively computable. We do this for three versions of NCC: a basic deterministic version, a probabilistic version and a version in which the computation can be subsidized by the center. The analysis turns out to depend on whether the inputs of the agents are probabilistically correlated or not and we analyze both cases.[1]

Keywords: Non-cooperative computation; Informational mechanism design; Cryptography

[*] Corresponding author.
E-mail address: shoham@stanford.edu (Y. Shoham).

[1] This work was supported by NSF Grant IIS-0205633.

0304-3975/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2005.05.009

1. Introduction

In this paper we introduce the framework of *non-cooperative computation* (*NCC*). In the NCC model, n agents each wish to compute an n-ary function w (we assume it is the same function for all agents, although that can be generalized), with each of the agents holding one of the inputs to w. For example, they may each hold a number and wish to compute the average. Or, to draw from the Boolean domain from which most of our technical results are drawn, they may each hold a bit (0 or 1) and wish to compute the majority function (which is 1 iff a majority of the input bits are). The process of computation is mediated by a center as follows: Each agent declares his input (truthfully or not) to the center, the center performs computation based on those inputs and reports back to the agents an output. In the basic setting we will define, the center simply applies w to the declared inputs and announces the value to all the agents. In two extensions of the basic setting we will consider the center is given greater flexibility, but in all cases the center's protocol is common knowledge among the agents.

The only thing standing between the agents and successful computation are their conflicting self interests. The incentives of the agents in NCC are multi-faceted, but are always defined in terms of the information available to the various agents. In this paper we will concentrate on agents whose utility function has two components. The first, called *correctness*, is the wish to compute the function correctly. The second, called *exclusivity*, is the wish that other agents do not compute the function correctly. We assume a lexicographic ordering between these two, with correctness preceding exclusivity.

As an example of this two-tiered preference ordering, imagine several biologists wishing to sequence a genome of an organism, each having deciphered a different part of the genome. Each of them would like to know the entire sequence so that s/he can publish a paper with the correct genetic code, but, given that, s/he would just as soon be a sole author. Assuming the scientists communicate via a center as described, the question is whether the scientists can be incented to reveal the correct code segments and thus all scientists will know the entire genetic code at the end. If the answer is yes, we will say that the function which assembles the entire genetic code from the individual segments is *non-cooperatively computable*, or NCC.[2]

To get a more technical intuition for this problem, let us consider again the Boolean domain. Specifically, consider agents $1, \ldots, n$ trying to compute some Boolean function $w(x_1, \ldots, x_n)$ where x_i is Boolean and known only by agent *i*. For example, consider the parity function (whose value is 1 iff the number of 1's in the input is even). Intuitively speaking, the parity function is not NCC; assuming all agents other than *i* disclose their true values, agent *i* has the incentive to lie; it will then reverse the result of the computation and obtain the correct value of w, whereas the others will end up with the wrong answer. In contrast, consider the majority function. Again, intuitively speaking, this function *is* non-cooperatively computable; if an agent attempts to deceive the others he will not in general be able to reconstruct the correct value himself.

These simple examples make it clear that the NCC framework is inherently game theoretic. Essential to the above arguments is the notion of equilibrium; we ask what an agent's

[2] By slight linguistic abuse, we use the abbreviation NCC as both a noun and an adjective.

best action is, given that the others adopt the equilibrium strategy (in this case, telling the truth). Indeed, NCC falls squarely in the area of *mechanism design* (or *implementation theory*) [5]. A branch of game theory that has attracted some attention in computer science recently, mechanism design is concerned with crafting protocols for self-interested agents that cause these agents behave in a certain desired way. What makes NCC unique from the standpoint of mechanism design is that the objective of the mechanism designer (the 'implemented function', to use the game theoretic jargon), as well as the utility functions of the individual agents, are defined entirely in terms of the information available to the different agents.[3]

It is instructive to contrast the NCC setting with the setting traditionally studied in cryptography, in particular the work on secure multi-party protocols (see [2] for a relatively recent overview and [4] for a discussion and overview of such protocols in a game-theoretic context). As in NCC, here too the goal is to compute a function jointly by a set of agents, each of whom holds part of the input. Furthermore, these agents are self-interested and even adversarial. The similarities end there, however. In the MPP literature there is an assumption that some of the agents (the 'good' agents) follow the prescribed protocol and the rest (the 'bad' agents) deviate from it. There are two models of deviation (the 'curious' and the 'malicious'), but the details do not concern us here. The key is that, with these assumptions, the traditional cryptographic setting involves no equilibrium analysis and indeed no explicit representation of the agents' utility functions; the latter are left implicit. It is an interesting exercise to attempt a game theoretic model of cryptographic protocols, as was done for the case of *byzantine agreement* [7]; such analysis exposes the non-comparable concerns of cryptography and game theory. In the discussion section at the end we comment on potential connections between NCC and the notion of *variable influence*, which is related to cryptography, but otherwise discussion of cryptography is beyond the scope of this article.

In order to state our specific results we must make several distinctions. The first distinction, familiar from the auction theory literature as well as from several computer science contexts, has to do with the information structure of the agents: Are their private inputs (signals, in the game-theory parlance) independent or correlated? The second distinction, very familiar in computer science, is between deterministic computation and probabilistic computation. The third important distinction, which is novel in computer science but standard in game theory, is whether the system supports the transfer of money and if so whether the mechanism is required to be budget balanced. (In plain terms, the question is whether the center can influence the behavior of the agents by injecting a subsidy into the system.)

[3] NCC is in fact a specialization of the more general category which we call *informational mechanism design*, or IMD. Recall that in general, any mechanism-design problem takes as input a social-choice function and the individual preferences of the players. IMD specializes MD by insisting that both the social-choice function and the individual preferences are purely informational; that is, they are defined in terms of which agent knows what information. NCC further specializes IMD by having a particular social-choice function; in NCC the desired outcome is that all agents know the value of the function w. That still leaves a key degree of freedom, namely the preferences of the players. In this paper, when we speak of NCC we implicitly assume the two-tiered lexicographic preference mentioned. But the concept of NCC is broader. For example, in follow-up work [6], the setting is augmented to capture other potential interests of agents: An agent may prefer that others not know its own input (so-called *privacy*) and an agent may prefer to know the inputs of other agents (so-called *voyeurism*). However, the basic properties of NCC are revealed already in the case considered here.

These last two restrictions give rise to three variants of NCC called, respectively, D-NCC (deterministic NCC), P-NCC (probabilistic NCC, in which the center given freedom to randomize its computation) and S-NCC (subsidized NCC). In the next section we provide the formal model of these and in the subsequent section we prove the following results:

1. In the independent values setting:
   (a) A Boolean function is D-NCC iff it is not dominated and not reversible.
   (b) A Boolean function is P-NCC iff it is D-NCC.
   (c) A Boolean function is S-NCC iff it is not reversible.
2. In the correlated values setting:
   (a) A Boolean function is D-NCC iff it is not dominated and not reversible.
   (b) A Boolean function is P-NCC iff it is not dominated.
   (c) Every Boolean function is S-NCC.

In addition, in the discussion section we go beyond Boolean functions and briefly discuss the NCC of k-order statistics.

2. Definitions

In this section we formally define the notion of NCC in the two-tiered preference setting. We first define the basic, deterministic case and then we define extensions of it.

2.1. Deterministic NCC (D-NCC)

Given a set of agents $N = \{1, 2, \dots, n\}$ and a special agent termed 'the center', we assume that there exists a private secure communication line between every agent $i \in N$ and the center. The type v_i of agent i is selected from some domain B_i. Although our definitions can be generalized to apply more broadly, our technical results primarily address the Boolean case, in which $B_i = B = \{0, 1\}$; from here on we will assume this restriction.

The vector of agent types $v = (v_1, \dots, v_n)$ is selected from a joint probability distribution p. We assume full support, i.e., $p(v) > 0$ for every $v \in B^n$. The function p induces functions p_i; for each $i \in N$ and $v_i \in B$, $p_i(v_i)$ is the marginal probability that agent i has type v_i. We say that we have an *independent values* setting if for every v we have that $p(v) = \Pi_{i=1}^{n} p_i(v_i)$. We say that we have a (*strictly*) *correlated values* setting if for every agent i, there exists $b_{-i} \in B^{n-1}$ such that $p(v_{-i} = b_{-i} | v_i = 0) \neq p(v_{-i} = b_{-i} | v_i = 1)$, where $v_{-i} = (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$. In intuitive terms, in the independent values setting the type of an agent does not tell it anything about the types of others, while in a correlated values setting it does.

Given a function $w : B^n \rightarrow B$, we consider the following protocol:

1. For any instantiated type vector $v \in B^n$, each agent i declares his type $\hat{v}_i$ to the center (truthfully or not; $\hat{v}_i = v_i$ may or may not hold).
2. The center computes the value $w(\hat{v}) = w(\hat{v}_1, \dots, \hat{v}_n)$ and announces it to all agents.
3. Each agent i computes $w(v)$ based on $w(\hat{v})$ and v_i (his true input).

The protocol defines a strategy space for each agent. A pure strategy for agent i is a pair of functions (f_i, g_i). $f_i : B \rightarrow B$, the *declaration function*, determines the input declared to the center based on the true input. Of particular interest will be the *truthful* declaration function, namely the identity function $f^t(v) = v$. $g_i : B^2 \rightarrow B$, the *interpretation function*, is used by the agent to decide on the value of the function based on the announcement by the center and his true input. Of particular interest will be the *trusting* interpretation function, namely the projection function $g^t(v_1, v_2) = v_1$ in which the agent simply accepts the value announced by the center. We will call the strategy (f^t, g^t) *straightforward*.

Note that the strategy profile[4] consisting only of straightforward strategies results in each agent's computing w correctly for all input vectors. We are interested in functions for which such a strategy profile forms an equilibrium. In this equilibrium, for each agent the straightforward strategy is a best response to all other agents' adopting the straightforward strategy. Of course, whether a strategy is a best response depends on the agent's preferences. The definition below captures the lexicographic ordering in each agent's preference, with correctness preceding exclusivity.

Throughout this paper, as we did for v_{-i} above, for any vector $(x_1, \ldots, x_n)$ we define $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. We also use (x_i, x_{-i}) to denote the reconstituted vector $(x_1, \ldots, x_n)$. For simplicity, we will often use $z(x_i, x_{-i})$ to denote the application of a function z to the vector (x_i, x_{-i}), rather than the more cumbersome $z((x_i, x_{-i}))$. The following definition applies to any domain B, though again we are concentrating on the Boolean domain $B = \{0, 1\}$.

Definition 1. Let N, p, w be as above. Then w is *deterministically non-cooperatively computable*[5], or D-NCC, if the following holds: For any agent i, every strategy (f_i, g_i) and every $v_i \in B$, it is the case that:

- either there exists $v_{-i} \in B_{-i}$ such that $g_i(w(f_i(v_i), v_{-i}), v_i) \neq w(v_i, v_{-i})$,
- or else for every $v_{-i} \in B_{-i}$ we have $w(f_i(v_i), v_{-i}) = w(v_i, v_{-i})$.

Note that the definition assumes that agents only care whether all other agents compute correctly or whether at least one of them does not. The definition takes no stance on whether the agent distinguishes among outcomes in which different nonempty sets of agents miscompute or among outcomes in which the agent itself miscomputes.

Two final comments about D-NCC. In D-NCC there is no discretion in designing the center's part of the protocol and so mechanism design reduces to equilibrium verification. In the more elaborate versions—P-NCC and S-NCC—there will be more discretion in this regard. Also, note that we have not considered more elaborate interaction among the players and the center, beyond this simple two-phase communication. Full discussion of this point lies outside the scope of the article and the reader can take this as an arbitrary restriction. However, the reader familiar with the *revelation principle* [5] will recognize that in fact no generality is lost by restricting the attention to this class of protocols.

2.2. From D-NCC to P-NCC

We will define two extensions of the basic D-NCC setting. In this section we discuss a probabilistic extension, termed P-NCC. This is a natural extension from the computer

[4] A strategy profile is vector of strategies, one for each agent.

[5] Under lexicographic ordering of correctness and exclusivity; we omit this comment in future definitions.

science perspective; in the next section we discuss a different extension that is natural from the game theoretic perspective.

In P-NCC we still look for equilibria in which agents adopt the straightforward strategy, but we allow the center—with some probability—to announce to the agents an incorrect value. We now have greater flexibility in deciding the protocol for the center. Rather than compute w, the center will compute n functions h_i, one for each agent (that is, given a declared type vector $\hat{v}$, the center will announce to agent i the value $h_i(\hat{v})$). In general we will have that $h_i \neq w$. Furthermore, in general h_i will be probabilistic; that is, $h_i : B^n \rightarrow \Delta(B)$, where $\Delta(B)$ is the set of probability distributions over B.

For $j \neq i$, define

$$E_j(i, f_i, g_i) = \Sigma_{(v_i, v_{-i}) \in B^n}[p(v_i, v_{-i})\text{Prob}(w(v) \neq h_j(f_i(v_i), v_{-i}))]$$

as the probability of agent j ending up with the wrong value of w, assuming all agents other than i follow the truthful equilibrium protocol, while i's strategy is (f_i, g_i). Note that in this expression, the expectation is taken both over p (the joint probability distribution over the inputs) and over *Prob* (which is determined by the probabilistic function h_j).

Of course, in the expression above E_j does not depend on g_i. However, we use this notation so that we can overload it and apply it when $i = j$ as well. This case, which captures i's estimation of his own probability of error, is defined by:

$$E_i(i, f_i, g_i) = \Sigma_{(v_i, v_{-i}) \in B}[p(v_i, v_{-i})\text{Prob}(w(v_i, v_{-i}) \neq g_i(h_i(f_i(v_i), v_{-i}), v_i))].$$

In the following definition, let f^t and g^t again denote the identity and projection functions, respectively.

Definition 2. Let N, p, w and E_j be as above. Let $0 < \delta \leqslant 1$. Then w is *probabilistically non-cooperatively computable with accuracy* δ, or δ-P-NCC, if there exist h_i such that the following both hold:

- For every $i \in N$ and $v \in B^n$, $\text{Prob}[w(v) \neq h_i(v)] < \delta$.
- For any agent i and any strategy (f_i, g_i),
  - either $E_i(i, f_i, g_i) > E_i(i, f^t, g^t)$,
  - or else the following are both true:
    - $E_i(i, f_i, g_i) = E_i(i, f^t, g^t)$ and
    - $E_j(i, f_i, g_i) \leqslant E_j(i, f^t, g^t)$ for some agent $j \neq i$.

In other words, again the straightforward strategy is a best response to the other agents' adopting the straightforward strategy; deviating either increases one's own probability of error, or leaves it unchanged while not increasing the error probability of others. In addition, we require a δ upper bound on the probability of error.

The non-relativized notion of P-NCC is then defined as follows:

Definition 3. Let N, p, w be as above. Then w is *probabilistically non-cooperatively computable*, or P-NCC, if w is δ-P-NCC for any $0 < \delta \leqslant 1$.

2.3. From D-NCC to S-NCC

So far we have assumed that the agents only derive utility from computing the function or denying others that benefit. We now add another ingredient to the mix, namely money. As in D-NCC we require that the center compute the function correctly; here again there is no discretion in this regard. However, in addition we give the center the power to allocate money to the agents by way of injecting additional incentives into the system. If the probabilistic extension in the previous subsection is natural in computer science, this extension is natural in economics and game theory. This subsidized variant of NCC is termed S-NCC.

Specifically, in S-NCC, as in D-NCC, the center is restricted to computing $w(\hat{v})$ and announcing the result to the agents. But in addition, the center has n (commonly known) payment functions $m_i : B^n \to \Re$; $m_i(\hat{v})$ is the payment from the center to agent i as a function of the declared values by all agents. In general, the payment $m_i(\hat{v})$ can be positive or negative and be of arbitrary magnitude. However, we will be interested in mechanisms in which in expectation the payment is positive and small.

Agent i has an error cost, which we normalize to be 0 if the computation is correct and 1 otherwise. The overall utility function of an agent who receives payment m and suffers error cost d is simply $m - d$. This so-called *quasi-linear utility function* [5] might raise two potential questions. First, one might wonder why it is reasonable to normalize different agents' utilities on the same scale. The answer is that in our game theoretic analysis, inter-agent comparison of utilities is not meaningful and does not impact equilibrium analysis. Second, one might ask why it is meaningful to correlate the scales of the error cost with the scale of money. The answer is that in some circumstances this could indeed be an important issue, but in the equilibria we will identify the amount of money is arbitrarily small and dominated by the error.[6]

Note that in S-NCC the strategy space of agents is larger than in D-NCC (or P-NCC). Specifically, any interpretation function g_i is now $g_i : B \times \Re \times B \to B$; the additional second argument is the payment received by the center. We extend the notion of the *trusting* interpretation function to the S-NCC setting in a natural way; continuing to use g^t to denote the trusting interpretation function, we define $g^t(r, x, v) = r$. The space of declaration functions remains unchanged in S-NCC and we continue to denote the truthful declaration function by f^t. Finally, we continue to use the term *straightforward strategy* for (f^t, g^t) as in the D-NCC and P-NCC settings.

For convenience, we will slightly overload the m_i function.

Given an input vector v and i's declaration strategy f_i, let

$$m_j(i, v, f_i) = m_j(f_i(v_i), v_{-i})$$

[6] An alternative definition, which would obviate these questions, would be to keep the monetary payoff separate and extend the two-tiered lexicographic preference structure to a three-tiered one, with correctness and monetary payoff both preceding exclusivity (the ordering between monetary payoff and correctness would be unimportant for our purposes). All our results would still hold under this model. However, it is convenient to amalgamate the first two into a direct overall utility function and the quasi-linear model is both natural and commonly used.

be the payment to agent *j*, assuming all other agents, excluding *i*, declare truthfully. Then, for any $v_i \in B$, define

$$Em_j(i, v_i, f_i) = \Sigma_{x \in B^n}[p(x \mid x_i = v_i) m_j(i, x, f_i)]$$

as the expected payment to agent *j* under the same conditions, conditional on agent *i*'s input being v_i.

Next we define $d_j(i, v, f_i, g_i)$ to be the error cost to agent *j* when the input vector is v and when all agents but *i* play the straightforward strategy (f^t, g^t) while *i* plays the pure strategy (f_i, g_i):

$$\begin{aligned} \text{for } j \neq i\colon d_j(i, v, f_i, g_i) &= 0 \quad \text{if } w(v) = w(f_i(v_i), v_{-i}),\\ &= 1 \quad \text{otherwise,}\\ \text{for } j = i\colon d_j(i, v, f_i, g_i) &= 0 \quad \text{if } w(v) = g_i(w(f_i(v_i), v_{-i}), m_i(i, v, f_i), v_i),\\ &= 1 \quad \text{otherwise.} \end{aligned}$$

For any i, j, the *direct utility* for *j* given the input vector and *i*'s strategy is given by

$$u_j(i, v, f_i, g_i) = m_j(i, v, f_i) - d_j(i, v, f_i, g_i)$$

and the expected version, conditional on *i*'s value by

$$Eu_j(i, v_i, f_i, g_i) = Em_j(i, v_i, f_i) - Ed_j(i, v, f_i, g_i).$$

With these definitions we define S-NCC as follows.

Definition 4. Let *N*, *p* and w be as above. Let $\varepsilon > 0$. Then w is *non-cooperatively computable with subsidy* ε, or ε-S-NCC, if there exist payment functions m_i as above for which the following holds for any agent *i*, any strategy (f_i, g_i) of *i* and every $v_i \in B$:

- either $Eu_i(i, v_i, f_i, g_i) < Eu_i(i, v_i, f^t, g^t)$ or else the following are both true:
  - ○ $Eu_i(i, v_i, f_i, g_i) = Eu_i(i, v_i, f^t, g^t)$ and
  - ○ $Ed_j(i, v_i, f_i, g_i) \geqslant Ed_j(i, v_i, f^t, g^t)$ for some agent $j \neq i$.
- $0 \leqslant Em_j(i, v, f^t) \leqslant \varepsilon$ for every $i, j \in N$ and every $v \in B^n$.

As in previous definitions, this one requires that it is an equilibrium for all agents to adopt the straightforward strategy. Note that this definition assumes that an agent cares about his error cost and his monetary payoff, as well as the error cost of the other agents, but not about the payments to the other agents.

Finally, analogously to the case of P-NCC, we define the non-relativized version of S-NCC:

Definition 5. Let *N*, *p* and w be as above. A function w is *subsidized non-cooperatively computable*, or S-NCC, if w is ε-S-NCC for every $\varepsilon > 0$.

3. Results for Boolean functions

Our goal in this section is to precisely characterize the Boolean NCC functions. We will provide six characterizations—of functions that are D-NCC, P-NCC and S-NCC, each for both the independent values case and the correlated values one.

We will need the following notions.

- A function $w : B^n \to B$ is called (*conditionally*) *dominated* if there exist an agent *i* and a $v_i \in B$ such that
  1. for all $y_{-i}, z_{-i} \in B^{n-1}$, it is the case that $w(v_i, y_{-i}) = w(v_i, z_{-i})$; and
  2. v_i is *relevant*, in that there exists $y_{-i} \in B^{n-1}$ such that $w(v_i, y_{-i}) \neq w(1-v_i, y_{-i})$.
- A function w is *reversible* if there exist $i \in N$ and $v_i \in B$ such that for every $y_{-i} \in B^{n-1}$ it is the case that $w(v_i, y_{-i}) = 1 - w(1 - v_i, y_{-i})$.[7]

Note: These definitions are reminiscent of, but different from, definitions in the social choice literature (e.g., [8]). We comment on this further in the discussion section.

3.1. Independent values

Our basic result concerns D-NCC:

Theorem 1. *In an independent values setting, a Boolean function is D-NCC iff it is not reversible and not dominated.*

Proof. Assume that the function w is reversible. Then there exists an agent *i* such that $w(0, y_{-i}) = 1 - w(1, y_{-i})$, for all $y_{-i} \in B^{n-1}$. Suppose all agents but *i* employ the straightforward strategy (f^t, g^t). Then (f^t, g^t) is not a best response for *i*. A better response is (f_i, g_i) where $f_i(v_i) = 1 - v_i$ and $g_i(y, v_i) = 1 - y$, for all $v_i, y \in B$ (this better response ensures that other agents always miscompute w but *i* never does). This proves that if w is reversible then it is not D-NCC.

Next assume that the function w is dominated. Then, there exist an agent *i* and $v_i \in B$ such that $w(v_i, y_{-i}) = w(v_i, z_{-i}) = d$ for every $y_{-i}, z_{-i} \in B^{n-1}$ and there exists $x_{-i} \in B^{n-1}$ such that $w(v_i, x_{-i}) \neq w(1 - v_i, x_{-i})$. Suppose all agents but *i* employ the straightforward strategy (f^t, g^t). Then again (f^t, g^t) is not a best response for *i*. A better response is (f_i, g_i) which differs from (f^t, g^t) only in that $f_i(v_i) = 1 - v_i$ and $g_i(y, v_i) = d$ for all $y \in B$ (this better response ensures that other agents miscompute $w(v_i, x_i)$ but *i* does not). This proves that if w is dominated then it is not D-NCC.

Finally, assume that the function w is neither reversible nor dominated. Consider agent *i* with strategy (f_i, g_i) and suppose all agents but *i* employ the straightforward strategy (f^t, g^t). Clearly, if *i* is irrelevant to w—that is, if $w(0, y_{-i}) = w(1, y_{-i})$ for all $y_{-i} \in B^{n-1}$—then $(f_i, g_i) = (f^t, g^t)$ is a best response for *i*. So assume that *i* is relevant and assume further that $f_i \neq f^t$. Suppose agent *i* has the true input v_i and declares value $f_i(v_i) = 1 - v_i$ and the center announces the value *r*. What could the value of $g_i(r, v_i)$ be? Since w is not dominated and since *i* is relevant, it cannot be that $g_i(r, v_i) = r$ without causing *i* to miscalculate for some inputs of the others. But at the same time it cannot be that $g_i(r, v_i) = 1 - r$, since this would imply that w is reversible. From this contradiction

[7] Note that for Boolean functions it is the case that if this property holds for v_i then it also holds for $1 - v_i$. Also note that among the *symmetric* functions, only the parity function (whose value is 1 iff an even number of its arguments are 1) and its negation are reversible, but there are many other non-symmetric reversible functions.

it follows that necessarily $f_i = f^t$. But clearly if $f_i = f^t$ then $(f_i, g_i) = (f^t, g^t)$ is a best response for *i* (if all agents including *i* declare truthfully, *i* only loses by deviating from the trusting interpretation function). The proves that if w is neither reversible nor dominated then it is D-NCC. □

P-NCC was introduced with the hope of increasing the power of NCC. The next result is disappointing in this respect, at least for the independent values context (but see the results for correlated values below):

Theorem 2. *In an independent values setting, a Boolean function is P-NCC if and only if it is D-NCC.*

Proof. Consider a Boolean function w. Trivially, if w is D-NCC then it is also P-NCC. Since a function is D-NCC iff it is neither dominated nor reversible, it is sufficient to show that if w is either dominated or reversible then it is not P-NCC.

Assume that w is dominated. Then, there exist $i \in N$, $d, v_i \in B$, such that for all $y_{-i}, z_{-i} \in B^{n-1}$ and some $x_{-i} \in B^{n-1}$ it is the case that $w(v_i, y_{-i}) = w(v_i, z_{-i}) = d$, while $w(1-v_i, x_{-i}) = 1-d$. Assume that w is P-NCC. Let $0 < \delta < 1$; then w is δ-P-NCC. This means that there exist functions h_i $(i = 1..n)$ for the center such that for each instance $v \in B^n$, $\text{Prob}[h_i(v) \neq w(v)] < \delta$. Assume that all agents $j \neq i$ play the straightforward strategy (f^t, g^t); it is enough to show that the straightforward strategy is not a best response for *i*. Consider the following strategy (f_i, g_i): $f_i(v_i) = f_i(1 - v_i) = 1 - v_i$ and for both $r \in B$, $g_i(r, v_i) = d$ while $g_i(r, 1 - v_i) = r$. We will show that (f_i, g_i) is a better response than (f^t, g^t). Clearly $E_i(i, f_i, g_i) \leqslant E_i(i, f^t, g^t)$. If $E_i(i, f_i, g_i) < E_i(i, f^t, g^t)$ then we are done; (f_i, g_i) is a better response for *i* than (f^t, g^t). So suppose $E_i(i, f_i, g_i) = E_i(i, f^t, g^t)$. To show that (f^t, g^t) is not a best response in this case either, we need to show that $E_j(i, f_i, g_i) > E_j(i, f^t, g^t)$ for some $j \neq i$. But now consider x_{-i} above, for which $w(1 - v_i, x_{-i}) = 1 - d$. Pick any agent $j \neq i$. Let $\text{Prob}[h_j(1 - v_i, x_{-i}) = d] = q$. For the straightforward strategy profile to be an equilibrium, it would have to be that $q < \delta$ (or else agent *j* would err with probability greater than δ on $w(1 - v_i, x_{-i})$, contradicting the definition of δ-P-NCC). Now consider $E_j(i, f_i, g_i)$ and define $r = p(v_i, x_{-i})$, the probability of the specific input vector (v_i, x_{-i}). It must be the case that $E_j(i, f_i, g_i) \geqslant r(1-q)$, since (a) $E_j(i, f_i, g_i)$ must be at least *r* times the probability that the center will announce to *j* the wrong value for this specific input (recall that according to f_i, for this input *i* announces $1 - v_i$) and (b) since the center announces d with probability q, it announces $1 - d$ with probability $1 - q$. And thus Since $q < \delta$, we have that $E_j(i, f_i, g_i) \geqslant r(1 - \delta)$. However, for $\delta < \frac{r}{1+r}$ we have that $r(1 - \delta) > \delta$ and so for small enough δ we have that $E_j(i, f_i, g_i) > \delta > E_j(i, f^t, g^t)$. This concludes the proof that if w is dominated then it is not P-NCC.

Now assume that w is reversible. Then, there exist $i \in N$, $v_i \in B$, such that $w(v_i, z_{-i}) = 1 - w(1 - v_i, z_{-i})$ for every $z_{-i} \in B^{n-1}$. Assume that w is P-NCC. Let $0 < \delta < 1$; then w is δ-P-NCC. Again, this means that there exist functions h_i for the center such that for each instance $v \in B^n$, $\text{Prob}(h_i(v) \neq w(v)) < \delta$. Assume that all agents $j \neq i$ play the straightforward strategy (f^t, g^t); it is enough to show that the straightforward strategy is not a best response for *i*. Let $\mu_1 = \Sigma_{z_{-i}\in B^{n-1}} p(z_{-i})\text{Prob}(w(v_i, z_{-i}) \neq h_i(v_i, z_{-i}))$ and $\mu_2 = \Sigma_{z_{-i}\in B^{n-1}}[p(z_{-i})(\text{Prob}(w(1-v_i, z_{-i}) \neq h_i(1-v_i, z_{-i})))]$.[8]

If $\mu_1 > \mu_2$, consider a deviation by *i* from the straightforward strategy to (f_i, g_i) where $f_i(v_i) = f_i(1-v_i) = 1-v_i$ and for both $r \in B$, $g_i(r, v_i) = 1-r$ while $g_i(r, 1-v_i) = r$. We get that $E_i(i, f_i, g_i) < E_i(i, f^t, g^t)$ and therefore the straightforward strategy is not a best-response.

If $\mu_2 > \mu_1$, consider a deviation by *i* from the straightforward strategy to (f_i, g_i) where $f_i(v_i) = f_i(1-v_i) = v_i$ and for both $r \in B$, $g_i(r, v_i) = r$ while $g_i(r, 1-v_i) = 1-r$. We get that $E_i(i, f_i, g_i) < E_i(i, f^t, g^t)$ and therefore again the straightforward strategy in not a best-response.

Finally, if $\mu_1 = \mu_2$, consider a deviation by agent *i* as in the case in which $\mu_1 > \mu_2$ above, i.e., a deviation to (f_i, g_i) where $f_i(v_i) = f_i(1-v_i) = 1-v_i$ and $g_i(r, v_i) = 1-r$, $g_i(r, 1-v_i) = r$, for any $r \in B$. We get that $E_i(i, f_i, g_i) = E_i(i, f^t, g^t)$. The proof now proceeds as in the proof for the dominated case (the following is an abridged version of that proof; the full proof can be substituted in here verbatim). To show that (f^t, g^t) is not a best response in this case either, we need to show that $E_j(i, f_i, g_i) > E_j(i, f^t, g^t)$ for some $j \neq i$. Consider $x_{-i} \in B^{n-1}$ that satisfies $d = w(v_i, x_{-i})$ and $1-d = w(1-v_i, x_{-i})$. Let $\text{Prob}[h_j(1-v_i, x_{-i}) = d] = q$. For the straightforward strategy profile to be an equilibrium, it would have to be that $q < \delta$ (or else agent *j* would err with probability greater than δ on $w(1-v_i, x_{-i})$, contradicting the definition of δ-P-NCC). Now consider $E_j(i, f_i, g_i)$ and again define $r = p(v_i, x_{-i})$. As before, it must be the case that $E_j(i, f_i, g_i) \geqslant r(1-q)$. But since $q < \delta$, we have that $E_j(i, f_i, g_i) \geqslant r(1-\delta)$ and thus for small enough δ we have that $E_j(i, f_i, g_i) > \delta > E_j(i, f^t, g^t)$. This concludes the proof that if w is reversible then it is not P-NCC. □

Finally, we consider the power of subsidies:

Theorem 3. *In an independent values setting, a function is S-NCC if and only if it is not reversible.*

Proof. From Theorem 1 we know that a function is D-NCC iff it is not dominated and not reversible. Since any function that is D-NCC is also S-NCC, it is enough to show that (a) dominated, non-reversible functions are always S-NCC and (b) reversible functions are never S-NCC.

The constant functions are non-dominated and non-reversible and therefore S-NCC (and even D-NCC), which is consistent with our theorem and therefore in the remainder of the proof we will consider only non-constant functions.

Let w be any non-constant Boolean function. We will first show that if w is dominated and non-reversible then it is S-NCC. That is, we show that for any given $\varepsilon > 0$, there exists payment functions m_j $(j = 1..n)$ such that for every agent *i*, if all agents $j \neq i$ play the straightforward strategy (f^t, g^t), then (f^t, g^t) is the best response for *i*; and that furthermore under these payment functions the expected payment to any agent is bounded by ε. We will prove this by construction. If agent *i* is irrelevant or there is no $v_i \in B$ such

[8] Note that in the independent values setting, $p(z_{-i}) = p(z_{-i} \mid v_i)$.

that $w(v_i, z_{-i}) = w(v_i, y_{-i})$ for all $y_{-i}, z_{-i} \in B^{n-1}$[9] then we take the payment to *i* to be identically 0. Since w is also non-reversible the best response for agent *i* is to use (f^t, g^t); this is exactly the proof as for the D-NCC case. Now consider an agent *i* with type $v_i \in B$, such that $w(v_i, z_{-i}) = d$ for every $z_{-i} \in B^{n-1}$. Given that w is not a constant function and is not reversible then it must be the case that there exist $y_{-i}, z_{-i} \in B^{n-1}$ such that $d = w(1 - v_i, y_{-i}) \neq (1 - v_i, z_{-i}) = 1 - d$. We determine the payment for agent *i* by $m_i(v_i, z_{-i}) = \delta$ and $m_i(1-v_i, z_{-i}) = 0$, for every $z_{-i} \in B^{n-1}$, where $0 < \delta < \varepsilon$ as will be determined below. Suppose that agent *i* deviates to (f_i, g_i) and that his type is $1 - v_i$. Then, there exist $y_{-i}, x_{-i} \in B^{n-1}$ such that $w(1 - v_i, y_{-i}) = d$ and $w(1 - v_i, x_{-i}) = 1 - d$. If $f_i(1-v_i) = 1-v_i$, then if $g_i(r, 0, 1-v_i) = 1-r$ for some $r \in B$ then $Eu_i(i, 1-v_i, f_i, g_i) < 0 = Eu_i(i, 1 - v_i, f^t, g^t)$ since such deviation will only cause miscomputation by *i* with non-zero probability. If $f_i(1 - v_i) = v_i$ then the output received from the center will be d and the payment to the agent is δ. Let $r_1 = p(1 - v_i, y_{-i})$ and let $r_2 = p(1 - v_i, x_{-i})$. If $g_i(d, \delta, 1 - v_i) = d$ then agent *i*'s computed output is wrong with probability of at least r_2 and if $g_i(d, \delta, 1 - v_i) = 1 - d$ then its output is wrong with probability of at least r_1. By taking $\delta < \min(r_1, r_2)$ we get that $Eu_i(i, 1 - v_i, f_i, g_i) \leqslant \delta - \min(r_1, r_2) < 0 = Eu_i(i, 1 - v_i, f^t, g^t)$ whenever $f_i(1 - v_i) = v_i$ and for every g_i. Together we get $Eu_i(i, 1 - v_i, f_i, g_i) < Eu_i(i, 1 - v, f^t, g^t)$ for every $(f_i, g_i) \neq (f^t, g^t)$. Now assume that agent *i*'s type is v_i. If $f_i(v_i) = 1-v_i$ then $Eu_i(i, v_i, f_i, g_i) < \delta = Eu_i(i, v_i, f^t, g^t)$. If $f_i(v_i) = v_i$ and $g_i(d, \delta, v_i) = 1-d$ then $Eu_i(i, v_i, f_i, g_i) \leqslant \delta-1 < \delta = Eu_i(i, v_i, f^t, g^t)$. Together we get $Eu_i(i, v_i, f_i, g_i) < Eu_i(i, v_i, f^t, g^t)$ for every $(f_i, g_i) \neq (f^t, g^t)$.

Given the above and since $\delta < \varepsilon$ we get that w is ε-S-NCC. Since this is true for any $\varepsilon > 0$, this proves that if any dominated, non-reversible function is S-NCC.

We now show that any reversible function is not S-NCC. Assume that w is reversible. In this case, there exist an agent *i* and type v_i of agent *i*, such that $w(v_i, y_{-i}) = 1 - w(1 - v_i, y_{-i})$ for every $y_{-i} \in B^{n-1}$. Assume that all agents, potentially excluding *i*, use the straightforward strategy (f^t, g^t). Let $\mu_0 = Eu_i(i, 0, f^t, g^t)$ and let $\mu_1 = Eu_i(i, 1, f^t, g^t)$. If $\mu_0 > \mu_1$ then deviating to (f_i, g_i), where $f_i(0) = f_i(1) = 0, g_i(0, t, 0) = 0, g_i(1, t, 0) = 1, g_i(0, t, 1) = 1, g_i(1, t, 1) = 0$ (for every *t*) satisfies $Eu_i(i, 1, f_i, g_i) > Eu_i(i, 1, f^t, g^t)$. Similarly, if $\mu_0 < \mu_1$ then deviating to (f_i, g_i), where $f_i(0) = f_i(1) = 1, g_i(0, t, 0) = 1, g_i(1, t, 0) = 0, g_i(0, t, 1) = 0, g_i(1, t, 1) = 1$ (for every *t*) satisfies $Eu_i(i, 0, f_i, g_i) > Eu_i(i, 0, f^t, g^t)$. Thus it is enough to consider the case in which $\mu_0 = \mu_1$. Let us consider deviation to (f_i, g_i), where $f_i(0) = f_i(1) = 0, g_i(0, t, 0) = 0, g_i(1, t, 0) = 1, g_i(0, t, 1) = 1, g_i(1, t, 1) = 0$ (for every *t*). In this case $Eu_i(i, v_i, f^t, g^t) = Eu_i(i, v_i, f_i, g_i)$ for every $v_i \in B$. However, agent $j \neq i$ will then compute a wrong answer whenever agent *i*'s type is 1, which happens with some positive probability. Hence, $Ed_j(i, 1, f_i, g_i) > 0 = Ed_j(i, 1, f^t, g^t)$ and w is not S-NCC. □

3.2. *Correlated values*

Intuitively speaking, in a correlated values setting we would expect more functions to be NCC than in the independent values case, since more information is conveyed by the private

[9] Note that this case is not precluded by the fact that w is dominated; it simply cannot be that *all* agents have these properties.

information. Imagine that the values to agents are assigned as follows: With probability $\frac{p}{2}$ all agents are assigned 1, with probability $\frac{p}{2}$ all agents are assigned 0 and with probability $1-p$ each agent's type is independently and uniformly selected from B. Imagine furthermore that p is large, for example 98%. Now if a given agent has the private value 1 he knows that with high probability the other agents do as well and thus can predict with high degree of accuracy the value of the function.

It is straightforward to see, given that in the correlated values setting there is still non-zero probability of every vector of types, that the set of D-NCC functions remains unchanged in the correlated values case:

Theorem 4. *In a correlated value setting, a Boolean function is D-NCC iff it is not dominated and not reversible.*

The proof is identical to the proof in the independent values case and is omitted.

In the remaining cases, however, correlated values do yield greater computing power. For reasons that will become clear, we skip P-NCC for the moment and speak about S-NCC. We have the following theorem:

Theorem 5. *In a correlated values setting, any Boolean function is S-NCC.*

Proof. Consider an arbitrary $\varepsilon > 0$. We will show that any Boolean function is ε-S-NCC in the correlated values case. By the definition of correlated values, we have that $\mathrm{Prob}[v_{-i} = y_{-i}|v_i = 0] \neq \mathrm{Prob}[v_{-i} = y_{-i}|v_i = 1]$ for some $y_{-i} \in B^{n-1}$. Let $p_{i,l} = \min(\mathrm{Prob}[v_{-i} = y_{-i}|v_i = 1], \mathrm{Prob}[v_{-i} = y_{-i}|v_i = 0])$, $p_{i,h} = \max(\mathrm{Prob}[v_{-i} = y_{-i}|v_i = 1], \mathrm{Prob}[v_{-i} = y_{-i}|v_i = 0])$. Let $v_{i,h}$ and $v_{i,l}$ be the types of agent i corresponding to $p_{i,h}$ and $p_{i,l}$, respectively.

The proof is again by construction and we set the payment functions as follows. An agent who announces $v_{i,l}$ gets nothing: $m_i(v_{i,l}, z_{-i}) = 0$ for every $z_{-i} \in B^{n-1}$. However, an agent who announces v_{i_h} gets a lottery, whose value is positive only under truthful declaration: $m_i(v_{i,h}, y_{-i}) = \varepsilon_i + 1$ (recall that y_{-i} is fixed here) and $m_i(v_{i,h}, z_{-i}) = \varepsilon_i - \frac{p_{i,h}}{1-p_{i,h}}$ for every $z_{-i} \neq y_{-i}$, where $0 < \varepsilon_i < \varepsilon$ and $\varepsilon_i < (1-p_{i,l})\left(\frac{p_{i,h}}{1-p_{i,h}}\right) - p_{i,l}$. Observe that $(1-p_{i,l})\left(\frac{p_{i,h}}{1-p_{i,h}}\right) - p_{i,l} > 0$.

If all agents use the straightforward strategy (f^t, g^t) then the payment to agent i with type $v_{i,l}$ is 0 and the expected payment to agent i with type $v_{i,h}$ is $p_{i,h}(\varepsilon_i + 1) + (1 - p_{i,h})\left(\varepsilon_i - \frac{p_{i,h}}{1-p_{i,h}}\right) = \varepsilon_i < \varepsilon$, as required. Consider a deviation by agent i from (f^t, g^t) to (f_i, g_i). Since $Eu_i(i, v, f^t, g_i) \leqslant Eu_i(i, v, f^t, g^t)$ and $Ed_j(i, v, f^t, g_i) = Ed_j(i, v, f^t, g^t)$ for every $v \in B$ and $j \neq i$ (such deviation will not change the payment to i and the fact the others will compute correctly, but might only make i compute incorrectly), it is enough to consider deviations where $f_i \neq f^t$. If agent i submits $v_{i,l}$ while his type is $v_{i,h}$ then he will be paid nothing instead of getting an expected payment of ε_i (and computing with no error) if he were to use the straightforward strategy. Therefore, $Eu_i(i, v_{i,h}, f_i, g_i) < Eu_i(i, v_{i,h}, f^t, g^t)$. Conversely, if agent i submits $v_{i,h}$ when his type his $v_{i,l}$ then his expected payment is $\varepsilon_i + p_{i,l} - (1-p_{i,l})\frac{p_{i,h}}{1-p_{i,h}}$. The latter however is negative, since $p_{i,l} - (1 - p_{i,l})\frac{p_{i,h}}{1-p_{i,h}} < p_{i,l} - p_{i,h} < 0$ and since we have selected $\varepsilon_i < (1 - p_{i,l})\left(\frac{p_{i,h}}{1-p_{i,h}}\right) - p_{i,l}$. This implies (since with (f^t, g^t) we get accurate computation and the above payments) that $Eu_i(i, v_{i,l}, f_i, g_i) < Eu_i(i, v_{i,l}, f^t, g^t)$.

From this it follows that w is ε-S-NCC. Furthermore, this is true for any $\varepsilon > 0$ and therefore w is S-NCC. □

We now turn to the remaining case, that of P-NCC. The positive result for S-NCC inspires us to look for a similar mechanism for the center, where the power of randomization compensates for the lack of monetary incentives. It turns out that the ability to randomize is not quite as powerful as the ability to print money, but it is not without power:

Theorem 6. *In a correlated values setting, a Boolean function is P-NCC if and only if it is not dominated.*

Proof. The proof that dominated functions are not P-NCC as in the independent values case; the correlation among the values plays no role in this direction.

For the other direction, assume that a function is not dominated. Let $\varepsilon > 0$. We will show that w is δ-P-NCC. It is sufficient to show that w is δ-P-NCC for sufficiently small δ, in particular for $\delta < 0.5$. We will make use of the following definitions:

Given a Boolean function w we distinguish between three types of agents:

1. Agent i is a *reverser* if for $v \in B$ we have that $w(v_i, z_{-i}) = 1 - w(1 - v_i, z_{-i})$ for every $z_{-i} \in B^{n-1}$.
2. Agent i is *irrelevant* if for every $v \in B, z_{-i} \in B^{n-1}$ we have that $w(v, z_{-i}) = w(1 - v, z_{-i})$
3. Agent i is *simple* if it is not a reverser and not irrelevant.

By the definition of correlated values, we have that $\text{Prob}[v_{-i} = y_{-i} | v_i = 0] \neq \text{Prob}[v_{-i} = y_{-i} | v_i = 1]$ for some $y_{-i} \in B^{n-1}$. Let $p_{i,l} = \min(\text{Prob}[v_{-i} = y_{-i} | v_i = 1], \text{Prob}[v_{-i} = y_{-i} | v_i = 0])$, $p_{i,h} = \max(\text{Prob}[v_{-i} = y_{-i} | v_i = 1], \text{Prob}[v_{-i} = y_{-i} | v_i = 0])$. Let $v_{i,h}$ and $v_{i,l}$ be the types of agent i corresponding to $p_{i,h}$ and $p_{i,l}$, respectively.

Let q_i be a random variable that gets the value γ_i if y_{-i} is declared by the agents in N_{-i} and $-\frac{p_{i,h}}{1-p_{i,h}}\gamma_i$ otherwise, where $\gamma_i > 0$ satisfies that $\max\left(\gamma_i, \frac{p_{i,h}}{1-p_{i,h}}\gamma_i\right) < \frac{\delta}{3}$. The expected value of q_i is 0 if agent i has the type $v_{i,h}$ and is $\gamma_i\left(p_{i,l} - (1 - p_{i,l})\frac{p_{i,h}}{1-p_{i,h}}\right) < \gamma_i(p_{i,l} - p_{i,h}) < 0$ if i has type $v_{i,l}$. Pick $\delta_i > 0$ such that $0 < \delta_i < \min\left(\gamma_i(p_{i,h} - p_{i,l}), \frac{\delta}{3}\right)$.

We now construct the functions h_i as follows. For any declaration vector $\hat{v}$, the center announces to agent i the value $w(\hat{v})$ with probability s_i and $1 - w(\hat{v})$ with probability $1 - s_i$, where s_i is determined as follows:

- If i is simple, $s_i = 1$.
- Otherwise (i is not simple), let the declared types be $\hat{v} = (\hat{v}_1, \ldots, \hat{v}_n)$. If $\hat{v}_i = v_{i,h}$ then $s_i = 1 - \frac{2}{3}\delta + \delta_i + q_i$ where the value of q_i is determined based on whether the other agents declared y_{-i} or not. Otherwise (if $\hat{v}_i = v_{i,l}$), $s_i = 1 - \frac{2}{3}\delta$. Observe that our selection of parameters satisfy that $0.5 < 1 - \delta < s_i < 1$.

Now consider agent i, with type $v_i \in B$ and the potential deviations of it to $(f_i, g_i) \neq (f^t, g^t)$.

- When the agent is simple then since w is also not dominated, we get that $E_i(i, f_i, g_i) > E_i(i, f^t, g^t)$ for every $(f_i, g_i) \neq (f^t, g^t)$; the proof is identical to the proof that non-dominated non-reversible functions are D-NCC (and the inequality is strict since we are considering only relevant agents).
- In order to deal with non-simple agents, we introduce the following construct. For $i \in N, v_i \in B$ let $t_i(v_i, d) = \Sigma_{z_{-i}\in B^{n-1}} p(v_i, z_{-i})\text{Prob}(w(d, z_{-i}) \neq h_i(d, z_{-i}))$, i.e. $t_i(v_i, d)$ is the probability the center will announce to i the right answer *for the input it receives* when i declares d and his type is v_i. Note the subtle definition: t_i refers only to correctness relative to the declared values. However, both d and v_i are relevant to assessing this correctness: d determines i's declaration and v_i induces a probability over the remaining declarations, given the joint distribution p over the inputs and the fact that the remaining agents play the straightforward function.

  We now first show that for any non-simple agent i and $v_i \in B$ we have that $t_i(v_i, 1-v_i) < t_i(v_i, v_i)$. Assume that a non-simple agent i, who has the type $v_{i,l}$, declares $v_{i,h}$ instead. Then the center will announce to him the right answer with probability $1 - \frac{2}{3}\delta + \delta_i + \gamma_i\left(p_{i,l} - (1-p_{i,l})\frac{p_{i,h}}{1-p_{i,h}}\right) < 1 - \frac{2}{3}\delta + \delta_i - \gamma_i(p_{i,h} - p_{i,l}) < 1 - \frac{2}{3}\delta = t_i(v_{i,l}, v_{i,l})$. Assume that i has type $v_{i,h}$ but declares $v_{i,l}$. In this case the center will announce to i the right answer with probability $1 - \frac{2}{3}\delta$, but has he declared $v_{i,h}$ the center would have announced the right answer with probability $1 - \frac{2}{3}\delta + \delta_i + p_{i,h} - (1-p_{i,h})\frac{p_{i,h}}{1-p_{i,h}} = 1 - \frac{2}{3}\delta + \delta_i > 1 - \frac{2}{3}\delta$. Hence, we get that for any non-simple agent i and $v_i \in B$ we have that $t_i(v_i, 1-v_i) < t_i(v_i, v_i)$.

  Assume i is a non-simple agent who uses the strategy (f_i, g_i) instead of the straightforward strategy (f^t, g^t), while all other agents use the straightforward strategy:

  - ○ If i is irrelevant and since $t_i(v_i, d) > 0.5$ for every $v_i, d \in B$, then for every f_i we have that $E_i(i, f_i, g^t) < E_i(i, f_i, g_i)$ where $g_i \neq g^t$. Since $t_i(v_i, 1-v_i) < t_i(v_i, v_i)$ for every $v_i \in B$ we get that $E_i(i, f^t, g^t) < E_i(i, f_i, g^t) < E_i(i, f_i, g_i)$ for every $f_i \neq f^t$ and $g_i \neq g^t$. Hence we get that $E_i(i, f^t, g^t) < E_i(i, f_i, g_i)$ for every $(f_i, g_i) \neq (f^t, g^t)$ when agent i is irrelevant.
  - ○ If i is a reverser, then if it declares $1 - v_i$ when his type is v_i and is announced r by the center, then since $w(v_i, z_{-i}) = 1 - w(1-v_i, z_{-i})$ for every $z_{-i} \in B^{n-1}$ we should have $g_i(r, 1-v_i) = 1 - r$ in order that deviation to (f_i, g_i) would be potentially profitable (otherwise agent i will compute the right answer with probability less than 0.5). This implies that the probability of computing the right answer by i when he uses (f_i, g_i) where $f_i(v) = 1 - v_i$ and his type is v_i is at most $t_i(v_i, 1-v_i)$, while when following (f^t, g^t) he will compute the correct answer when his type is v_i with probability $t_i(v_i, v_i) > t_i(v_i, 1-v_i)$. Hence we get that $E_i(i, f^t, g^t) < E_i(i, f_i, g_i)$ for every $(f_i, g_i) \neq (f^t, g^t)$ when agent i is a reverser.

Since the probability of providing the right answer is always greater than $1 - \delta$ and for every agent i we have that $E_i(i, f^t, g^t) < E_i(i, f_i, g_i)$ for every $(f_i, g_i) \neq (f^t, g^t)$, we have that w is δ-P-NCC. Since the above construction is defined for any sufficiently small error probability $\delta > 0$, we get that w is P-NCC. □

4. Discussion

In this paper we introduced the concept of non-cooperative computing NCC and defined three flavors of it—deterministic (D-NCC), probabilistic (P-NCC) and subsidized (S-NCC). The NCC framework is very broad and one of our goals has been to simply put it on the research map. In addition, we provided a comprehensive analysis of the class of Boolean functions that are NCC when the utility function of agents is defined by correctness and exclusivity, ordered lexicographically. For this case our results are summarized in the following table (each cell in the table specifies the necessary and sufficient conditions for the function to be NCC in the corresponding setting):

	D-NCC	P-NCC	S-NCC
independent values	not reversible and not dominated	not reversible and not dominated	not reversible
correlated values	not reversible and not dominated	not dominated	*any*

We have restricted our results to Boolean functions. This was done in order to make our discussion more concrete, while concentrating on a class of functions that is central in computer science. Nevertheless, our definitions can be easily extended to more general domains and further results shown. For example, with appropriate extension of the definitions, it can be shown that in the independent values setting, the k-order statistic is D-NCC for $1 < k < n$, while the max and min functions are not D-NCC; however, the max and min functions are S-NCC. However, pursuing these extensions is beyond the scope of this article.

As we discussed in the introduction, the NCC framework is quite distinct from other frameworks and in particular from those encountered traditionally in cryptography. Let us nonetheless conclude with an open question regarding an interesting potential three-way connection between NCC, social-choice theory [8] and a specific notion related to cryptography, namely *variable influence* [3,4]. We will not repeat the definitions or results from these areas and so these comments will be meaningful particularly to the reader familiar with one or both of these fields. Indeed, such a reader will undoubtedly have noticed the surface similarity, as well as the deep differences. In particular, both social choice and variable influence appeal to the notion of dictatorship, which is stronger than our notion of (conditional) dominance. Conversely, the most elegant proofs of the seminal result in social choice theory—Arrow's impossibility theorem [1,9]—use the notion of a 'pivotal agent', which in some sense is a weaker notion than our notion of a reverser agent and the related notion of function reversibility. And so at this stage we can point to no crisp technical connections between NCC and either social choice theory or variable influence. By the same token, there are to date no established connections between social choice theory and variable influence, despite the fact that such connections were one of the motivations for studying variable influence.[10] These three pairwise connections seem to us to merit further investigation.

[10] N. Linial, personal communication.

Acknowledgements

We thank Felix Brandt, Samuel Ieong and Tim Roughgarden who graciously took time to read the manuscript and who provided many useful comments. We especially thank Ryan Porter for detailed comments and discussions. Of course all remaining errors are ours alone.

References

[1] K.J. Arrow, Social Choice and Individual Values, second ed., Yale University Press, New Haven, CI, 1963 (1st Ed. published 1951).

[2] O. Goldreich, Secure multi-party computation, Working draft, Weizmann Institute, 1998.

[3] J. Kahn, G. Kalai, N. Linial, The influence of variables on Boolean functions, in: Proc. 29th IEEE Symp. on Foundations of Computer Science, 1985, pp. 68–80.

[4] N. Linial, Game-theoretic aspects of computer science, in: R.J. Aumann, S. Hart (Eds.), Handbook of Game Theory with Economic Applications, Vol. II, North-Holland, Amsterdam, 1994, pp. 1340–1395, (Chapter 38).

[5] A. Mas-Colell, M.D. Whinston, J.R. Green, Microeconomic Theory, Oxford University Press, Oxford, 1995.

[6] R. McGrew, R.W. Porter, Y. Shoham, Towards a general theory of non-cooperative computing, in: Proc. Ninth Conf. on Theoretical Aspects of Rationality and Knowledge (TARK), 2003.

[7] S. Morris, H.S. Shin, Approximate common knowledge and co-ordination: recent lessons from game theory, J. Logic, Language Inform. 6 (1997) 171–190.

[8] H. Moulin, Social choice, in: R.J. Aumann, S. Hart (Eds.), Handbook of Game Theory with Economic Applications, Vol. II, North-Holland, Amsterdam, 1994, pp. 1091–1125 (Chapter 31).

[9] P. Reny, Arrow's theorem and the Gibbard–Satterthwaite theorem: a united approach, Econom. Lett. 70 (2001) 99–105.